The Evolution of Fluoroquinolone-Resistance in In Vitro and in Natural Populations of *Mycobacterium tuberculosis*

INAUGURALDISSERTATION
zur
Erlangung der Würde eines Doktors der Philosophie

vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von
Rhastin Allan Del Rio Castro
aus Kanada

Basel, 2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Herrn Prof. Dr. Sebastien Gagneux und Herrn Prof. Dr. Martin Ackermann

Basel, den 17 December 2019

Prof. Dr. Martin Spiess

Dekan

dedicated to my loving parents,
Allan and Ana

Contents

Acknowledgements

John Donne famously wrote that "no man is an Island." That phrase could not be any more true during the process of making this Thesis. Throughout the entirety of my research work, never once did I feel unsupported. I was always surrounded by people who cared; people who were either willing to give me a hand when I needed help, or lend me an ear when I needed advice. All of your support helped make these last four years such a fulfilling, formative, and memorable experience, and for that, I am truly grateful.

First and foremost, I would like to thank my PhD supervisor, Prof. Dr. Sebastien Gagneux. Within the first few months of joining your lab first as a Master's student six years ago, I knew I wanted to stay on for a PhD. From the very beginning, you made me feel like I belonged here and capable of doing the work. No amount of words could substitute how much confidence that simple act instilled in me. You always gave me guidance when I needed it, but also the space and time to pursue ideas and projects that I thought were promising. As I've mentioned before, your work encouraged me to come to Basel, but your enthusiasm and thoughtfulness was what convinced me to stay. I consider myself incredibly fortunate to have been able to work for you, and I sincerely thank you.

To Prof. Dr. Martin Ackermann, thank you very much for taking the time to serve as the co-referee of my PhD committee and for the evaluation of this Thesis. I want to give special thanks to Prof. Dr. Pascal Mäser as well for kindly chairing my PhD defense.

I would like to deeply thank Drs. Sonia Borrell and Andrej Trauner, who have also served as my mentors since I first joined the group. At every planning and research design step, experimental work, data analysis, manuscript writing, presentations preparations, and a whole slew of other tasks, you gave me incredible guidance and motivated me to improve and challenge myself. I would not know half of the things I know now, nor would I be capable of the skills I'm proud of being good at doing, without all of your teachings and support.

To all the past and current members of the Tuberculosis Research Unit at the Swiss TPH that I had the opportunity to work with, thank you for your constant support, advice, and lively conversations. I would like to especially acknowledge Ms. Miriam Reinhard and Ms. Julia Feldmann for providing me with the initial experimental training, continued support, and kind advice through all of these years; the experimental work presented in this Thesis was made possible with your help. To Mr. Lujeko Kamwela, thank you for being a fantastic student for me to mentor; the data you generated and analyzed for the mutational profiling assays was instrumental in pushing our manuscript into becoming something we were all incredibly proud of, and I sincerely thank you for your incredible effort and hard work. To Dr. Daniela Brites, thank you very much for your input on the genomics work, and for providing kind and inspiring advice on academia and life in general; I very much appreciated them. To Ms. Michaela Zwyer and Ms. Tatjana Meyer, thank you for all of the random but great conversations that really helped me

get through some tough working days; I always looked forward to our shared coffee breaks.

I want to especially thank the "PT" office, Dr. Sebastian Gygli, Dr. Liliana Rutaihwa, Ms. Monica Ticlla, and Ms. Chloé Loiseau, for the wonderful friendships and camaraderie. I'm both grateful and astounded at how well we complemented each other's humour and eccentricities. I bet it's not often that five people crammed into a space as small as our office for many years could still come out as good friends. "Believe me. We are, THE behhhst....." I will cherish all the weird reference jokes ("Gym? What's a gym?"), sudden yelling about code, hushed-tone conversations behind closed doors, random bursts of laughter, even more random single-noted whistling or humming, and all the other incredible memories we made. Although all of our shared PhD journeys are ending, wherever we go all go from here, please know you have a friend in me.

To Dr. Amanda Ross, thank you for the huge amount of support in developing the mathematical modeling work of this Thesis and for the great conversations; your efforts in reviewing our manuscript improved it immensely.

To the wonderful people I had the fortune to be friends with at the Swiss TPH, I would like to thank all of you for providing me with an amazing and lively environment to be in. Whether it was relaxing near the Rhein river, during a weekend party, in a hut situated on the top of a snow-covered mountains in the Alps, or simply during a quiet lunch or coffee somewhere near the TPH, your kindness and sincerity made me feel like I found a second home. I would like to thank the following people for helping me create an incredible PhD journey: Dr. Natalie Wiedemar, Ms. Anna Fesser, Dr. Astrid Knoblauch, Dr. Severine Erismann, Dr. Nerina Vischer, Dr. Henry Ntuku, Dr. Francis Mhimbira, Dr. Gordana Panic, Dr. Jana Kovac, Mr. Oliver Bärenbold, Dr. Elizaveta Semenova, Dr. Sabelo Dlamini, Dr. Simone Sutherland, Dr. Emilie Pothin, Ms. Marta Palmeirim, Ms. Nadja Wipf, Ms. Andrea Leuenberger, Dr. Wendelin Moser, Dr. Harris Héritier, Mr. Martin Matuska, Mr. Christos Kokaliaris, Mr. Anton Beloconi, Dr. Josephine Malinga, Dr. Laura Ruckstuhl, Dr. Katya Galactionova, Dr. Flavia Camponovo, Dr. Mari Dumbaugh, Dr. Isabel Zenklusen, Dr. Julian Rothen, Ms. Anneth Tumbo, Ms. Camilla Messerli, Ms. Kate Harlan, Dr. Jerry Hella, Dr. Prince Asare, Dr. Isaac Darko Otchere, Ms. Theresa Reiker, Dr. Herry Mpesi, Ms. Katrina Obas, Ms. Carla Grolimud, Ms. Monica Cal, Dr. Kirsten Gillingwater, Ms. Jennifer Giovanoli Evack, Mr. Halim Mahmoud, Mr. Jacek Korecki, Ms. Lara Courty, Dr. Machteld "Maggie" Wyss, Ms. Lena Koever, and Dr. Noëmi Meier.

I would also like to take this opportunity to remember the late Dr. Richard M'Bra, one of our former PhD student colleagues who was always so full of life and brought so much joy to those around him; you will be missed, my friend.

To Ms. Christine Mensch, I want to sincerely thank you for being the most welcoming person to myself and to the PhD student community at the Swiss TPH at large. Your support throughout my time at the Swiss TPH made me feel that I was always part of a great community.

To Prof. Dr. Nino Künzli, thank you for all of the initiative and support in helping us create the best possible environment to develop and study in. I would also like to thank you for supporting the

three large PhD student excursion events that we organized; I can assure you that it was a PhD career highlight for every individual involved.

To Ms. Anja Bührer, I want to sincerely and deeply thank you for inspiring me every day to become the best person I can be; you always believe in me, and helped me stand tall during times when I felt that I could not stand at all (sometimes, literally). I would also like to thank Leyla for the incredible amount of joy that she provides to our household, and for helping nurse me back to health whenever I was ill; you're a very good dog.

Last, but definitely not least, I want to thank my lovely parents, Mr. Ronald Allan Castro and Mrs. Ana Maria Castro. Twenty-three years ago, you gave up everything to move from your home country with very little money to a foreign place that could, at the time, only offer promises of a better life. You chased this promise not for yourselves, but to be able to provide the best possible future for my siblings and myself. While being college-educated in your native country, you first had to work in menial jobs just to make ends meet and provide for us. But you had a vision, and a plan. You worked hard, got another college education, fell in love with your adopted country, and after many years decided to call it home by buying a house you could have only dreamed of when you first started. Through all of this moving about and moving up, there was always one constant: your love and willingness to provide the best possible opportunities for us. I cannot express my gratitude enough for setting us up from something very little to the best possible life we could have asked for. I can only say that I put my absolute everything into the writing of this Thesis, and I wholeheartedly dedicate it to you. Thank you.

Summary

Despite the advent of antimicrobials, tuberculosis (TB) caused the greatest amounts of deaths due to an infectious disease in 2018, accounting for 1.2 million deaths alone and an additional 0.3 million deaths due to TB-HIV co-infections. Current TB treatment regimens may impose substantial economic and logistical burdens for patients and health care systems, as even treatment against drug-susceptible strains of *Mycobacterium tuberculosis* (*Mtb*), the aetological agent of TB, requires daily doses of antimicrobials for 6 to 9 months. Further complications arise when patients are infected with multidrug-resistant strains of *Mtb*, which increases treatment duration to 9 to 24 months. Treatment success rates are also reduced from approximately 85% for drug-susceptible cases down to almost 50% for multidrug-resistant cases of TB. Therefore, substantial efforts by the medical and research communities are currently underway to develop new treatment regimens that are both more efficacious and can reduce treatment duration against both drug-susceptible and multidrug-resistant strains of *Mtb*.

Fluoroquinolones (FQs) form a vital component in established and experimental TB treatment regimens. Older generations of FQs have been used to treat multidrug-resistant forms of TB, while newer and more potent FQs are being tested in experimental regimens that aim to reduce TB treatment duration. Extensive biochemical work has shown that FQs target DNA gyrase, the sole type II topoisomerase in *Mtb*. Molecular epidemiological studies demonstrate that clinically-relevant FQ-resistance (FQ-R) mutations are restricted to a limited set of chromosomal mutations in the genes encoding DNA gyrase: *gyrA* and *gyrB*. However, little work has been done on exploring the evolution of FQ-R in populations of *Mtb*. Investigating how different *Mtb* populations evolve under FQs pressure, and how FQ-R mutations affect the continued evolution of *Mtb* populations, may aid in maintaining the potency and potential use of FQs.

Treatment regimens for TB generally use standardized, empirical dosing, including when using FQs. Previous work has shown that different *Mtb* genetic variants can associate with different frequencies of drug-resistance (DR), even when using standardized treatment regimen. Bacterial genetics have also been shown to modulate the phenotypes that DR mutations confer. Whether the genetic variation present in natural populations of *Mtb* would also modulate the frequency of FQ-R emergence, or the phenotypes that FQ-R mutations confer, is currently not known. It is also unclear how FQ-R mutations themselves affect the continued evolution of *Mtb* populations.

In this Thesis, we explored how FQ-R evolves in *Mtb*. Specifically, we used extensive in vitro experiments coupled with analysis of publicly available *Mtb* genomic sequences isolated from clinical strains to test whether the genetic variation in *Mtb* modulated the frequency and phenotypes of FQ-R mutations. We then used a mathematical modeling framework and *in silico* simulations to test the relative contributions of bacterial factors hypothesized to be relevant in DR evolution in determining the frequency of FQ-R. Lastly, we used further in vitro assays and *Mtb* genomic sequences from clinical

strains to test the impact of FQ-R mutations on the continued evolution of *Mtb* populations.

This Thesis consists of 6 main chapters. The first chapter provides a general introduction into TB and the rationale for this Thesis, while the second chapter states the main Aims and Objectives. Three chapters then present the results of our research work, with one chapter dedicated per stated Objective. The last chapter provides a synopsis of our main findings, states general limitations, highlights the contribution of this Thesis to the *Mtb* and antimicrobial resistance research communities, and highlights potential future directions that can build upon this Thesis work.

In *Chapter 1*, we introduce the global burden of TB and treatment regimens for TB. We then highlight the problem of antimicrobial resistance, and the potential use of more potent FQs in reducing TB treatment times. Lastly, we introduce evolutionary concepts in DR evolution, highlight the role of bacterial genetics in DR evolution, and state the rationale for this Thesis.

In *Chapter 2*, we state the Main Aim and Objectives of this Thesis.

In *Chapter 3*, we use the Luria-Delbrück fluctuation analysis, further in vitro assays, and genomic sequences analysis to test whether the genetic variation present in natural populations of *Mtb* modulates the frequency and phenotypes of FQ-R mutations. We show that the *Mtb* genetic background can lead to differences in FQ-R that spans two orders of magnitude. Furthermore, we find that the *Mtb* genetic background modulates the phenotypes conferred by FQ-R mutations in vitro, and the observed types and relative frequencies of FQ-R mutations both in vitro and in the clinic.

In *Chapter 4*, we adapt a mathematical modeling framework developed by Ford et al., 2013 to simulate the frequency of FQ-R in order to test the relative contributions of different bacterial factors in FQ-R evolution. Our results suggest that not all relevant bacterial factors have been accounted for in the model, and that a new model of DR evolution is required for *Mtb*.

In *Chapter 5*, we again use the Luria-Delbrück fluctuation analysis coupled with analysis of *Mtb* genomes to test whether FQ-R mutations can affect the continued evolution of *Mtb* populations. We observe that FQ-R mutations can increase the frequency of acquiring further DR mutations in *Mtb* in vitro. However, genomic analysis demonstrate that FQ-R mutations do not necessarily associate with increased genetic diversity in natural populations of *Mtb*.

In *Chapter 6*, we highlight the key findings of this Thesis. We then state the limitations, discuss the implications of our results, and propose future directions in the study of FQ-R evolution in *Mtb*.

Abbreviations

AMR	Antimicrobial resistance
CFU	Colony forming units
CTAB	Cetyl trimethylammonium bromide
DNA	Deoxyribonucleic acid
DR	Drug-resistance
DS	Drug-susceptible
DSB(s)	Double-stranded DNA break(s)
DST	Drug-susceptibility testing
FQ(s)	Fluoroquinolone(s)
FQ-R	Fluoroquinolone-resistance
HGT	Horizontal gene transfer
HIV	Human immunodeficiency virus
INH	Isoniazid
INH-R	Isoniazid-resistance
L1, L2, ..., L7	Lineage 1, Lineage 2, ..., Lineage 7
L-J	Löwenstein-Jensen
LTBI	Latent tuberculosis infection
MDR-TB	Multidrug-resistant tuberculosis
MIC	Minimum inhibitory concentration
MSS-MLE	Ma, Sarkar, Sandri-Maximum Likelihood Estimator method
Mtb	*Mycobacterium tuberculosis*
MTBC	*Mycobacterium tuberculosis* complex
OFX	Ofloxacin
PCR	Polymerase chain reaction
QBP	Quinolone-binding pocket
QRDR	Quinolone-resistance-determining region
RIF	Rifampicin
RIF-R	Rifampicin-resistance
SNP(s)	Single nucleotide polymorphism(s)
STR	Streptomycin
STR-R	Streptomycin-resistance
TB	Tuberculosis
WHO	World Health Organization

WT	Wild-type
XDR-TB	Extensively drug-resistant tuberculosis

List of Figures

List of Tables

1 General Introduction

1.1 Global Burden of Tuberculosis

For millennia, tuberculosis (TB) has been a scourge on humanity. Morphological and molecular evidence suggests that TB was present in people that lived during prehistoric Eastern Mediterranean (Hershkovitz et al., 2008), prehistoric East Asia (Suzuki et al., 2008), in ancient Egypt (Nerlich et al., 1997), and in Pre-Columbian South America (Bos et al., 2014). Today, TB remains a global burden on human health and mortality, with approximately 10 million incident cases in 2018 (WHO, 2019a).

TB is a communicable disease. Transmission of TB occurs through the inhalation of aerosols infected with bacterial species belonging to the *Mycobacterium tuberculosis* complex (MTBC), the etiological agents of TB, which includes the human-adapted species *M. tuberculosis* (*Mtb*) and *M. africanum* (Koch, 1882; Gagneux, 2018). Once infected, most individuals control the infection, do not develop active disease, and are classified as having latent TB infection (LTBI) (O'Garra et al., 2013; Lin et al., 2018). However, individuals with LTBI have a 5-15% of progressing to active and symptomatic disease within 2 years of infection (O'Garra et al., 2013; Lin et al., 2018). Individuals who progress from initial infection or LTBI to active TB disease develop fevers, heavy night sweats, and weight loss (O'Garra et al., 2013; Lin et al., 2018); these patients also begin to cough out infected aerosols, thereby beginning the infection cycle anew (O'Garra et al., 2013; Lin et al., 2018). While multiple immunological mechanisms are hypothesized to determine whether and how individuals progress to active TB disease, immunosuppression, such as through HIV co-infection, is a strong risk factor (Ernst, 2012; O'Garra et al., 2013; Lin et al., 2018).

The burden of TB is not spread equally across the world (Figures 1.1 & 1.2). This is due to TB primarily being a disease of poverty. Poor living conditions, under-nutrition, overcrowding, poor sanitation, and poor indoor air quality are some well-known risk factors for TB (Oxlade et al., 2012; Dheda et al., 2016). People from developing nations are therefore at higher risk of being infected with MTBC, developing active TB, and then further transmitting MTBC. Indeed, while Europe and the Americas each accounted for approximately 3% of TB cases in 2018, South-East Asia accounted for 44%, Africa for 24%, and the Western Pacific for 18% (WHO, 2019a). Furthermore, eight countries (India, China, Indonesia, the Philippines, Pakistan, Nigeria, Bangladesh, and South Africa) accounted for approximately two-thirds of all TB cases (WHO, 2019a).

Because of the massive burden TB imposes, TB remains one of the top ten causes of death in humans (WHO, 2019a). In 2018, around 1.2 million deaths were due to TB alone, with approximately 0.25 million additional deaths due to TB-HIV co-infections (WHO, 2019a). Since 2007, TB has been the leading cause of deaths in humans due to a single infectious agent, surpassing HIV (WHO, 2019a).

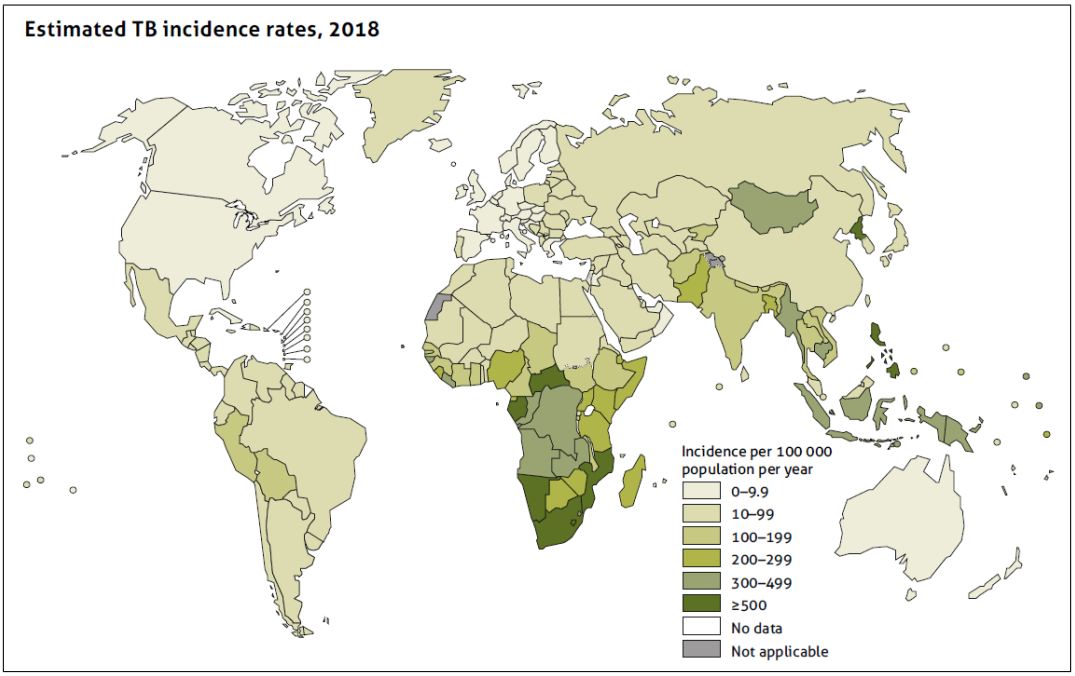


Figure 1.1: Estimated TB Incidence Rates in 2018.
From the Global Tuberculosis Report 2019, World Health Organization (WHO) (WHO, 2019a).

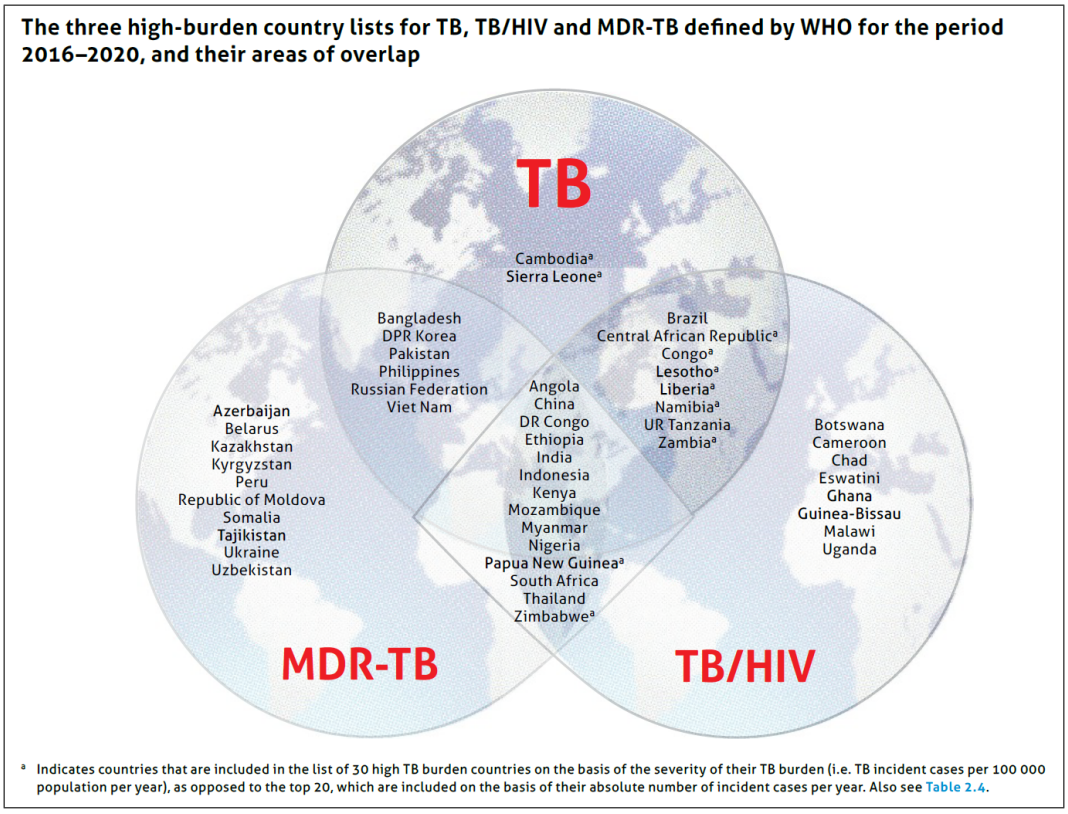


Figure 1.2: List of high-burden countries for TB, for TB/HIV co-infections, MDR-TB, and their overlaps for the period 2016-2020, as defined by the WHO.
From the Global Tuberculosis Report 2019, World Health Organization (WHO) (WHO, 2019a).

1.2 Treatment of Tuberculosis

TB-related mortality would likely even be higher if not for the advent of antimicrobials. While *Mtb* was discovered by Robert Koch in 1882 (Koch, 1882), the first antibiotic drugs with anti-TB activity, streptomycin and *para*-aminosalicylic acid, would only be discovered in 1944 and 1946, respectively (Schatz et al., 1944; Lehmann, 1946). The modern age of TB treatment regimen was then ushered in by the discovery of isoniazid (INH) in 1952 (Robitzek et al., 1952), rifampicin (RIF) in 1966 (Maggi et al., 1966), and their first combined use in 1977 (Dickinson et al., 1977). Today, INH and RIF continue to be used in combination with pyrazinamide (discovered in 1948; McKenzie et al., 1948) and ethambutol (discovered in 1961; Thomas et al., 1961) as the first-line treatment regimen against drug-susceptible *Mtb* (WHO, 2017; WHO, 2019a). This first-line regimen uses standardized, empirical dosing and has high efficiency in the clinic, with an approximately 85% treatment success rate (Farah et al., 2005; Bao et al., 2007; Gebrezgabiher et al., 2016; Tiberi et al., 2018).

However, the first-line regimen requires patients taking daily doses of antibiotics for at least 6 months (WHO, 2017). This substantial pill burden can lower patient adherence, and can impose large economic burdens on patients and health care systems (Munro et al., 2007; Barter et al., 2012; Alipanah et al., 2018; Ruru et al., 2018). Therefore, current research and development of new TB treatment regimens focus on developing new drugs or drug combinations that both improve treatment outcomes and reduce treatment duration and pill burdens (Tiberi et al., 2018; Vjecha et al., 2018).

1.3 Antimicrobial Resistance in Tuberculosis

The rise of antimicrobial resistance (AMR) further complicates TB treatment efforts. In general, AMR in pathogens represents a global health crisis as it increases treatment failures, treatment duration, treatment costs, and likelihood of adverse side effects from treatment (MacGowan, 2008; Winston et al., 2012; Laxminarayan et al., 2013; Kibret et al., 2017; Zhang et al., 2018). The prevalence of AMR is determined by the complex interaction between multiple behavioural, socioeconomic, health systems, and biological factors (Laxminarayan et al., 2013). For instance, poverty (Laxminarayan et al., 2013; Alvarez-Uria et al., 2016; Stosic et al., 2018), patient non-adherence (Laxminarayan et al., 2013; Stosic et al., 2018), improper or inadequate use of antibiotics (Laxminarayan et al., 2013; Shah et al., 2017; Alipanah et al., 2018), and lack of patient support (Laxminarayan et al., 2013; Alipanah et al., 2018) have all been positively associated with AMR prevalence.

AMR in TB is of particular importance; it represents the single largest cause of mortality due to AMR in pathogens, accounting for approximately 200,000 out of the nearly 700,000 AMR-related deaths in 2014 (O'Neill, 2016). The biggest risk factor for AMR in TB is history of previous treatment (Dalton et al., 2012; Zhao et al., 2012; Dean et al., 2017). Armed conflict and the collapse of established health systems have also been implicated in increased AMR prevalence in TB (Eldholm et al., 2016). In 2018, there was an estimated half a million cases of multidrug-resistant TB (MDR-TB), defined

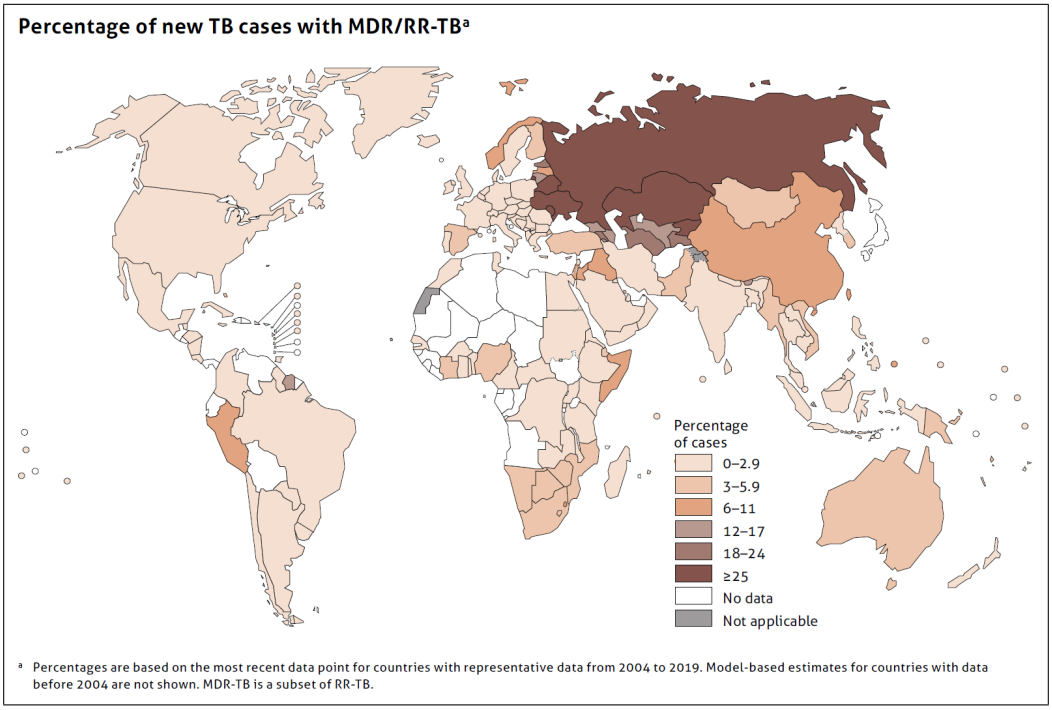


Figure 1.3: Estimated percentage of new TB cases that were MDR-TB or RIF-R TB in 2018. From the Global Tuberculosis Report 2019, World Health Organization (WHO). (WHO, 2019a)

as an infection with an MTBC strain that is resistant to at least RIF and INH (Figures 1.3, 1.4, & 1.5; WHO, 2019a). Treatment for MDR-TB presents a medical, economic, and logistical challenge. While patients with fully drug-susceptible MTBC strains require 6 to 9 months of treatment, patients with MDR-TB require 9 to 24 months. Furthermore, treatment success rates for MDR-TB are generally lower compared to drug-susceptible TB, with MDR-TB treatment success rates ranging from 79.8% in randomized and controlled clinical trials (Nunn et al., 2019) down to approximately 50.0% in observational studies (Kibret et al., 2017; Parmar et al., 2018; Zhang et al., 2018; WHO, 2019a). Thus, there is a need to develop new drugs or drug combinations that both improve treatment success rates and reduce treatment duration for drug-susceptible and drug-resistant cases of TB.

1.4 Use of Fluoroquinolones for Tuberculosis Treatment

New types of fluoroquinolones (FQs) are being tested in experimental treatment regimens against both drug-susceptible and drug-resistant cases of TB (Imperial et al., 2018; Vjecha et al., 2018). FQs are broad-spectrum antibiotics that have been and continue to be used for TB treatment (Aldred et al., 2014; WHO, 2019b). Previous MDR-TB treatment regimens used the second-generation FQs ciprofloxacin or ofloxacin, or the third-generation FQ levofloxacin (Yew et al., 2003; Chan et al., 2004; Takiff et al., 2011); today, current MDR-TB treatment regimens use the third-generation FQs moxifloxacin or gatifloxacin (Van Deun et al., 2010; Nunn et al., 2019; WHO, 2019b). However, moxifloxacin and gatifloxacin are also being considered as part of new treatment regimens that aim to

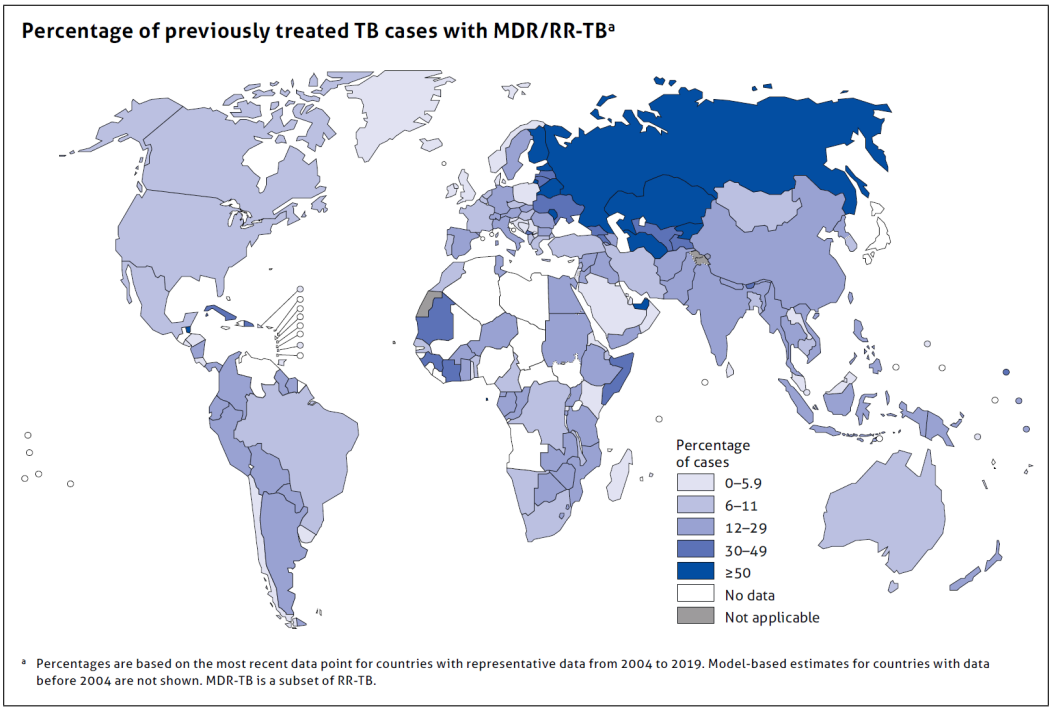


Figure 1.4: Estimated percentage of previously treated TB cases that were MDR-TB or RIF-R TB in 2018.
From the Global Tuberculosis Report 2019, World Health Organization (WHO) (WHO, 2019a).

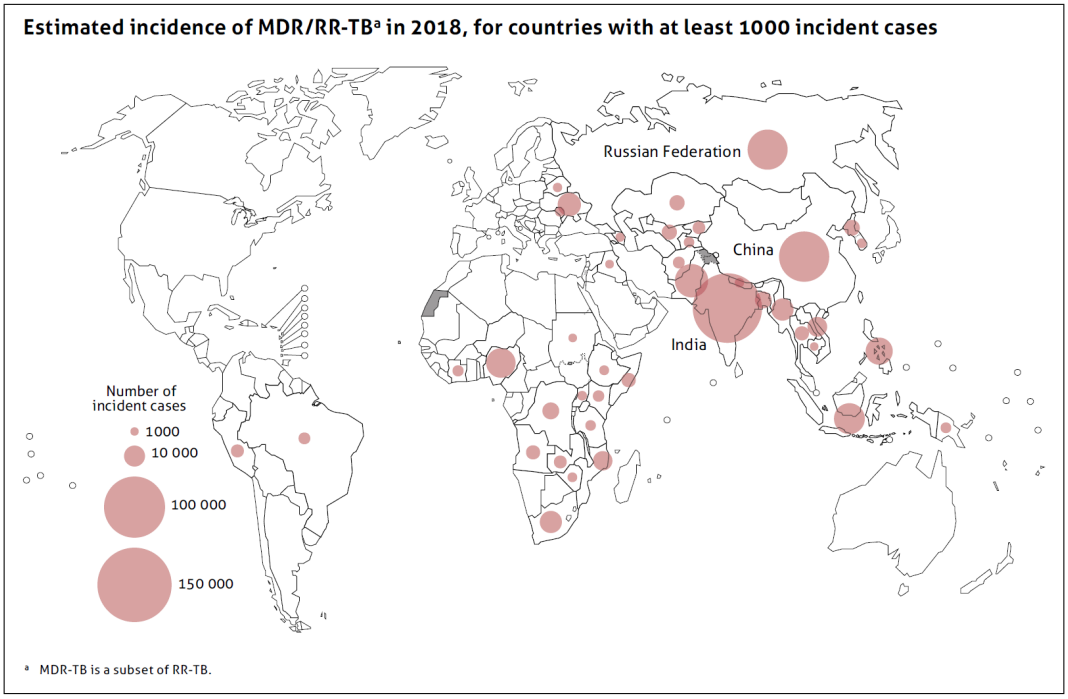


Figure 1.5: Estimated incidence of MDR-TB or RIF-R TB in 2018, for countries with at least 1000 incident cases.
From the Global Tuberculosis Report 2019, World Health Organization (WHO) (WHO, 2019a).

improve treatment success rates and reduce treatment duration for drug-susceptible cases of TB (Gillespie et al., 2014; Jindani et al., 2014; Merle et al., 2014; Imperial et al., 2018; Vjecha et al., 2018) and for MDR-TB (Van Deun et al., 2010; Nunn et al., 2019). These newer generations of FQs are more potent than their older counterparts (Ji et al., 1995; Ji et al., 1998; Yu et al., 2016), have been associated with greater treatment success rates in both *Mtb*-infected mice and humans (Nuermberger et al., 2004; Ahmad et al., 2018), and may even be used against MTBC strains that are resistant to the previous generations of FQs (Chien et al., 2016; Maitre et al., 2017; Pranger et al., 2019).

Population-based studies have shown regional and drug-specific differences in FQ-resistance (FQ-R) prevalence, with a range of 1.0–16.6% of all TB cases being ofloxacin-resistant, 0.5–12.4% being levofloxacin-resistant, and 0.9-14.6% being moxifloxacin-resistant (Zignol et al., 2016). Due to treatment practices, current FQ-R prevalence generally occurs in the context of MDR-TB (WHO, 2019a). Extensively drug-resistant TB (XDR-TB), defined as cases of MDR-TB that have additional FQ-R and resistance to an injectable aminoglycoside, represents one of the biggest challenge in TB treatment, with treatment success rates being approximately 30% (Leimane et al., 2010; Alene et al., 2017; WHO, 2019a). Thus, mitigating further increases in FQ-R prevalence is important to save patient lives, as well as to sustain the current and potential use of FQs. Understanding how FQ-R is acquired and maintained in natural populations of MTBC may provide insights in the development of new tools or strategies that restrict further increases in FQ-R prevalence.

In bacteria, FQs target the type II topoisomerases DNA gyrase and topoisomerase IV (Aldred et al., 2014; Mayer et al., 2014). Type II topoisomerases introduces negative supercoils in the bacterial chromosome through an ATP-dependent process that first creates a double-stranded DNA break (DSB) in the chromosome, loops another intact segment of the chromosome through the DSB, and then religates the DSB (Gellert et al., 1976; Levine et al., 1998). Type II topoisomerases are therefore essential in regulating DNA topology, DNA replication, and transcription (Levine et al., 1998). DNA gyrase is the only type II topoisomerase present in MTBC, and is therefore the sole target of FQs in MTBC (Cole et al., 1998; Mayer et al., 2014). Specifically, FQs bind to the active site of DNA gyrase, with the region designated as the "Quinolone-binding pocket" (QBP) (Piton et al., 2010; Aldred et al., 2016; Blower et al., 2016). Clinically-relevant FQ-R mutations in MTBC are restricted to a small subset of chromosomal mutations in the "quinolone-resistance-determining region" (QRDR) of the two genes that encode DNA gyrase: *gyrA* and *gyrB* (Takiff et al., 1994; Maruri et al., 2012; Wollenberg et al., 2017). Structural characterizations have shown that amino acid residues in the QRDR share the same physical space as the QBP, showing a direct relationship between FQ-DNA gyrase binding and FQ-R mutations (Piton et al., 2010; Aldred et al., 2016; Blower et al., 2016). Further, QRDR mutations have been shown to change the structure and interaction potential of DNA gyrase, including a disruption of the DNA gyrase-FQs interaction that effectively leads to the FQ-R phenotype (Piton et al., 2010; Aldred et al., 2016; Blower et al., 2016; Pandey et al., 2018).

While plenty of literature exists on the relevant biochemical mechanisms for FQ-R in MTBC,

the evolutionary dynamics of FQ-R in MTBC has been less studied. Indeed, although multiple factors contribute to the prevalence of AMR in general (Laxminarayan et al., 2013), the emergence of AMR within any given pathogen population is ultimately an evolutionary process (Wiesch et al., 2011; Hughes et al., 2017). Understanding how MTBC populations evolve under FQ pressure, as well as how MTBC populations continue in their evolution once they are FQ-resistant, may provide insights into how to mitigate further increases in FQ-R prevalence.

1.5 Population Biology Factors that modulate Antimicrobial Resistance Evolution

In general, the evolution of AMR in pathogens is modulated by the interaction between multiple biological factors. Firstly, the emergence of AMR mutations is positively associated with the rate of genetic diversity production (Wiesch et al., 2011; Hughes et al., 2017). Production of genetic diversity can be achieved through DNA replication errors (Reha-Krantz, 2010) or DNA repair mechanism-induced mutagenesis (Ysern et al., 1990; Gong et al., 2005; Baharoglu et al., 2014; Chapman et al., 2012), which together make up the DNA mutation rate (Wiesch et al., 2011; Hughes et al., 2017). Increased DNA mutation rates are positively associated with increased AMR prevalence in in vitro and in natural populations of multiple bacterial species (Oliver et al., 2000; Chopra et al., 2003; Örlén et al., 2006; Oliver et al., 2010; Torres-Barceló et al., 2013; Wielgoss et al., 2013; Wang et al., 2013; Couce et al., 2016). Horizontal gene transfer presents another avenue for increasing genetic diversity in bacterial populations, as different bacterial strains or even species can exchange unique genetic material (Hughes et al., 2017; Sun et al., 2019). Because these unique genetic material can contain AMR mutations, the rate of HGT has been positively associated with increased AMR prevalence (Hughes et al., 2017; Sun et al., 2019). However, unique genetic material must first be produced through DNA mutations; thus, the impact of HGT on AMR prevalence is also modulated by the DNA mutation rate (Sun et al., 2019).

Genetic diversity is positively associated with population size, as there is simply a greater probability of mutational events in a larger populations compared to smaller populations (Hughes et al., 2017; Frenoy et al., 2018); thus, population sizes are also hypothesized to be positively associated with AMR prevalence (Hughes et al., 2017; Frenoy et al., 2018).

The number of potential mutations that can confer the AMR phenotype, also known as the AMR mutational target size, may modulate AMR prevalence (Hughes et al., 2017). The AMR mutational target size is itself modulated by the drug type and drug concentration that the pathogen population is exposed to (Takiff et al., 2011; McGrath et al., 2014; Hughes et al., 2017). For instance, high drug concentrations would generally lead to a smaller mutational target size, as less AMR mutations would be capable of providing the high level of resistance required to maintain growth under these conditions (Lindsey et al., 2013; Ford et al., 2013; McGrath et al., 2014; Huseby et al., 2017; Hughes et al., 2017). Furthermore, different drugs or drug types may have differences in biochemical interactions with the target biomolecule, or have different target biomolecules altogether, which may lead to differences in

AMR mutational target sizes (Zhou et al., 2000; Takiff et al., 2011; Ford et al., 2013; McGrath et al., 2014; Hughes et al., 2017). Conditions that lead to large AMR mutational target sizes, such as the use of low drug concentration or drugs that have multiple biomolecule targets, are positively associated with AMR prevalence (Ford et al., 2013; McGrath et al., 2014; Gygli et al., 2017; Hughes et al., 2017).

While the AMR mutational target size determines the number of potential mutations that may be acquired, the fitness effect of the AMR mutations determines their relative frequencies and transmission potential (Wiesch et al., 2011; Hughes et al., 2017). Because antibiotics generally target essential and evolutionary conserved biomolecules or pathways, AMR mutations generally confer a fitness cost (Andersson et al., 2010; Hughes et al., 2017). However, some AMR mutations have been shown to confer little or no fitness costs in in vitro assays, and these AMR mutations are generally the most prevalent in the clinic (Gagneux et al., 2006c; Andersson et al., 2010; Gygli et al., 2017; Huseby et al., 2017; Wollenberg et al., 2017). The fitness effect of AMR mutations is thus hypothesized to modulate the clinical prevalence of AMR as well (Andersson et al., 2010; Hughes et al., 2017).

Non-heritable phenotypic mutations may also modulate AMR prevalence. Phenotypic mutations that alter protein structure and function occur due to protein promiscuity, or due to errors in transcription, translation, or epigenetic modification (Loftfield et al., 1972; Ozbudak et al., 2002; Payne et al., 2019b). Phenotypic mutations can lead to antimicrobial tolerance, antimicrobial persistence, or protein expression changes, which themselves have been positively associated with AMR prevalence (Balaban et al., 2004; Javid et al., 2014; Gygli et al., 2017; Levin-Reisman et al., 2017; Chaudhuri et al., 2018; Hicks et al., 2018).

In natural populations of MTBC, AMR evolution is driven primarily by the emergence and maintenance of chromosomal mutations (Gygli et al., 2017). Specifically, HGT- or plasmid-based resistance have not been documented in MTBC (Boritsch et al., 2016; Gygli et al., 2017). Furthermore, while the upregulation of efflux pumps have been observed in vitro, their relevance in the clinic is under debate (Gygli et al., 2017). Lastly, antimicrobial persistence and tolerance may potentiate the acquisition of AMR mutations in in vitro and in mice-infecting populations of MTBC (Javid et al., 2014; Chaudhuri et al., 2018); however, the only antimicrobial persistence and tolerance mechanisms that have been shown to associate with the clinical prevalence of AMR were themselves conferred by chromosomal mutations (Hicks et al., 2018). Therefore, from an evolutionary perspective, studying AMR evolution in MTBC, and FQ-R evolution specifically, provides a unique setting to investigate the emergence and maintenance of clinically-relevant chromosomal mutations that confer or lead to AMR.

1.6 Role of Bacterial Genetics in Antimicrobial Resistance Evolution

To study aspects in MTBC evolution, the genetic diversity present in natural populations of MTBC must be taken into account. Although MTBC genetic diversity is small compared to other bacterial pathogens, the global population of human-adapted MTBC species (*i.e. Mtb* and *M. africanum*) can be grouped into seven distinct genetic groups (Figure 1.6A; Comas et al., 2010; Gagneux, 2018);

these genetic groups, named Lineages 1 through 7, have non-random phylogeographic distributions (Figure 1.6B; Gagneux et al., 2006b; Gagneux, 2018). For example, while strains belonging to Lineage 1 are generally found along the rim of the Indian Ocean and in Southeast Asia (Douglas et al., 2003; Wan et al., 2017; Gagneux, 2018), Lineage 2 and Lineage 4 strains are found throughout the world, with Lineage 4 strains being the most prevalent globally (Gagneux et al., 2006b; Stucki et al., 2016; Gagneux, 2018). Different lineages have also been associated with differences in phenotypes, including differences in growth rates (Sarkar et al., 2012; Gehre et al., 2013), gene expression profiles (Rose et al., 2013), progression to active disease (Jong et al., 2008; Baya et al., 2019), and the induction of host immune cell responses (Portevin et al., 2011; Reiling et al., 2013).

Bacterial genetics may also play role in the evolution of FQ-R in MTBC. Indeed, bacterial genetics have been shown to modulate AMR evolution in general in MTBC (Gagneux, 2018). For instance, Lineage 2 strains have repeatedly been associated with MDR-TB in the clinic (Borrell et al., 2009; Casali et al., 2014; Merker et al., 2015; Eldholm et al., 2016); this association has been hypothesized to be due to Lineage 2 strains having much higher mutation rates, or a "hypermutator" phenotype, which consequently leads to a higher rate of AMR emergence in Lineage 2 strains compared to strains from other lineages (Rad et al., 2003; Ford et al., 2013). Bacterial genetics have also been shown to modulate the mutational target size for both INH-R (Gagneux et al., 2006a; Fenner et al., 2012) and RIF-R (Zaczek et al., 2009; Ford et al., 2013) in MTBC. Lastly, the MTBC genetic background has also been shown to modulate the fitness effects of RIF-R mutations (Gagneux et al., 2006c). Whether bacterial genetics modulate the rate of FQ-R emergence, the mutational target size for FQ-R, or the fitness effect of FQ-R mutations in MTBC have yet to be determined.

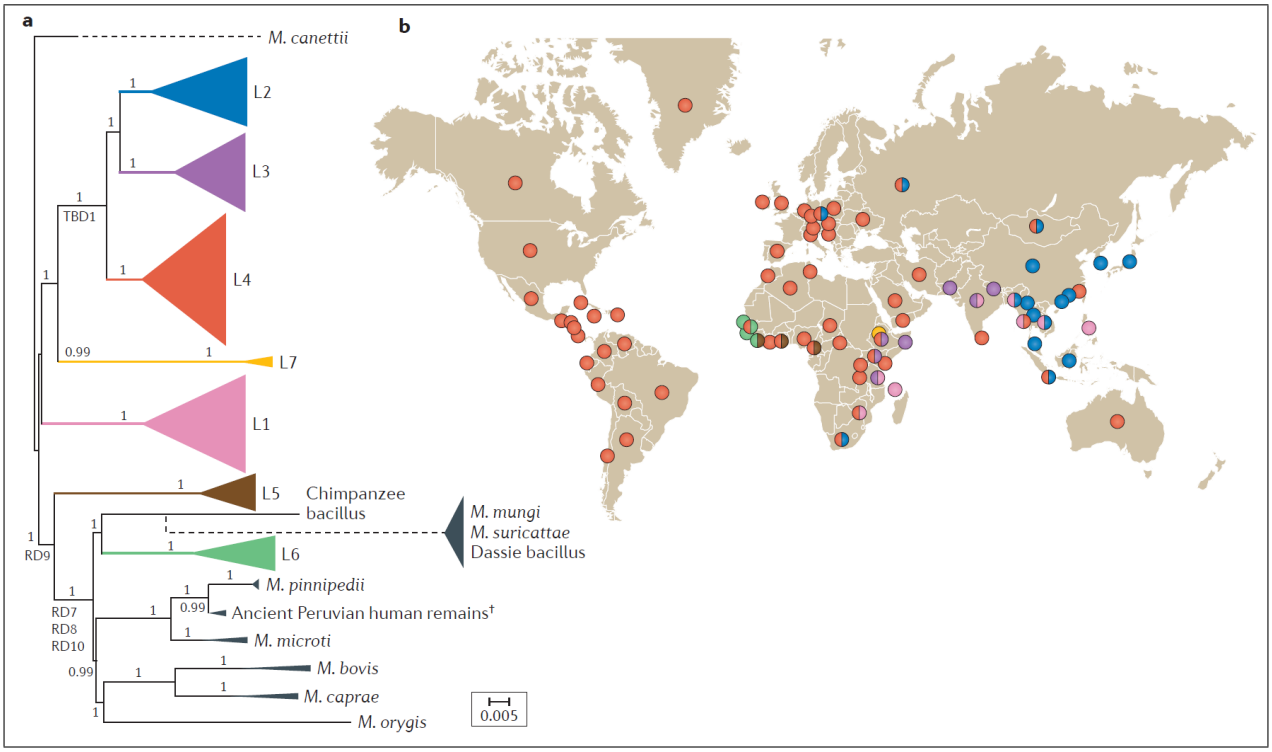


Figure 1.6: Phylogeny and Global Distribution of the *Mycobacterium tuberculosis* Complex. (A) Phylogeny of the *M. tuberculosis* Complex (MTBC) based on genomic data previously published by Bos et al., 2014; phylogeny is rooted using the *M. canettii* genome. Branches of the seven human-adapted lineages (in colour, named Lineages 1 through 7) and of the animal-adapted lineages (in grey) are represented as collapsed triangles for clarity. Lineage (L)1, L2, L3, L4 and L7 are the genetic groups belonging to *M. tuberculosis* sensu stricto, while L5 and L6 are the genetic groups belonging to *M. africanum*. TBD1 represents the shared genomic deletion between L2, L3, and L4, while shared deletions in the region of difference 7 (RD7), RD8, RD9, and RD10 are indicated under their respective branches. Dashed line connecting the *M. mungi*, *M. suricattae*, and dassie bacillus represents most likely phylogenetic relationship to the rest of the MTBC based on genetic data published in Dippenaar et al., 2015 and Alexander et al., 2016. Dagger on "Ancient Peruvian human remains" denotes an approximately 1,000-year-old MTBC DNA isolated from archaelogical human remains in Peru and genome sequenced by Bos et al., 2014. Scale bar represents number of nucleotide substitutions per year. Bootstrap confidence intervals are stated where applicable. (B) Global phylogeographic distribution of the seven human-adapted bacterial lineages of MTBC, with colours denoting lineage. Figure from Gagneux, 2018.

1.7 Rationale

For this Thesis, we investigated whether bacterial genetics can modulate the rate of FQ-R emergence, the mutational target size for FQ-R, and the fitness effect of FQ-R mutations in MTBC. We focused our exploration on whether the genetic background of *Mtb* sensu stricto could modulate FQ-R evolution; this is because genetic groups in *Mtb* include Lineages 1, 2, and 4, making *Mtb* responsible for the greatest TB disease burden amongst the MTBC (Comas et al., 2010; Gagneux, 2018). Understanding how FQ-R evolves in MTBC may provide insights in designing new treatment regimens that restricts the emergence and continued transmission of FQ-R in the clinic.

As mentioned previously, FQ-R mutations may lead to structural changes in DNA gyrase, an evolutionary-conserved enzyme that produces and religates DSBs to achieve its function of regulating DNA topology (Levine et al., 1998; Piton et al., 2010; Aldred et al., 2016; Blower et al., 2016). We hypothesize that FQ-R mutations-induced structural changes in DNA gyrase may reduce the efficiency of its DSB religation function, leading to increased expression of error-prone DNA repair mechanisms. Therefore, we also explored whether FQ-R mutations themselves can affect the evolutionary trajectory of MTBC populations by modulating DNA mutation rates and, consequently, modulating the levels of genetic diversity present.

2 Aims and Objectives

2.1 Aims

The main aim of this Thesis was to investigate the evolution of fluoroquinolone-resistance in *M. tuberculosis*. Specifically, this Thesis used in vitro experimental work, a mathematical modeling coupled with an *in silico* simulations framework, and analysis of whole-genome sequencing data isolated from clinical strains to explore whether the genetic variation present in natural *M. tuberculosis* populations modulates fluoroquinolone-resistance evolution. Furthermore, this Thesis tested whether the presence of fluoroquinolone-resistance mutations themselves can modulate the evolutionary potential of *M. tuberculosis* populations.

2.2 Objectives

2.2.1 Objective 1

Determine whether the genetic variation present in natural populations of *M. tuberculosis* influences the evolution of flouroquinolone-resistance (Chapter 3).

2.2.2 Objective 2

Use a mathematical model to explore the relative contributions of bacterial factors hypothesized to modulate the evolution of fluoroquinolone-resistance in *M. tuberculosis* (Chapter 4).

2.2.3 Objective 3

Investigate whether fluoroquinolone-resistance mutations associate with increased genetic diversity in *M. tuberculosis* populations (Chapter 5).

3 The Genetic Background modulates the Evolution of Fluoroquinolone-Resistance in *Mycobacterium tuberculosis*

Rhastin A.D. Castro[1,2], Amanda Ross[1,2], Lujeko Kamwela[1,2], Miriam Reinhard[1,2], Chloé Loiseau[1,2], Julia Feldmann[1,2], Sonia Borrell[1,2], Andrej Trauner[1,2,*], and Sebastien Gagneux[1,2,*]

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland

[2]University of Basel, Basel, Switzerland

[*]Corresponding Authors: E-mails: andrej.trauner@swisstph.ch; sebastien.gagneux@swisstph.ch.

This Chapter has been published in *Molecular Biology and Evolution* 2019, msz214, https://doi.org/10.1093/molbev/msz214

3.1 Abstract

Fluoroquinolones (FQ) form the backbone in experimental treatment regimens against drug-susceptible tuberculosis. However, little is known on whether the genetic variation present in natural populations of *Mycobacterium tuberculosis* affects the evolution of FQ-resistance (FQ-R). To investigate this question, we used nine genetically distinct drug susceptible clinical isolates of *Mtb* and measured their frequency of resistance to the FQ ofloxacin (OFX) in vitro. We found that the *Mtb* genetic background led to differences in the frequency of OFX-resistance (OFX-R) that spanned two orders of magnitude and substantially modulated the observed mutational profiles for OFX-R. Further, in vitro assays showed that the genetic background also influenced the minimum inhibitory concentration and the fitness effect conferred by a given OFX-R mutation. To test the clinical relevance of our in vitro work, we surveyed the mutational profile for FQ-R in publicly available genomic sequences from clinical *Mtb* isolates, and found substantial *Mtb* lineage dependent variability. Comparison of the clinical and the in vitro mutational profiles for FQ-R showed that 51% and 39% of the variability in the clinical frequency of FQ-R *gyrA* mutation events in Lineage 2 and Lineage 4 strains, respectively, can be attributed to how *Mtb* evolves FQ-R in vitro. As the *Mtb* genetic background strongly influenced the evolution of FQ-R in vitro, we conclude that the genetic background of *Mtb* also impacts the evolution of FQ-R in the clinic.

Key words: *Mycobacterium tuberculosis*, antimicrobial resistance, evolution, fluoroquinolones, epistasis, mycobacteria, fitness.

3.2 Introduction

Antimicrobial resistance (AMR) poses a major threat to our ability to treat infectious diseases (MacGowan, 2008; Winston et al., 2012). The rise of AMR is a complex phenomenon with a broad range of contributing socioeconomic and behavioral factors (Dalton et al., 2012; Merker et al., 2015; Alvarez-Uria et al., 2016; Eldholm et al., 2016; Shah et al., 2017). However, the emergence of AMR within any pathogen population is ultimately an evolutionary process (Wiesch et al., 2011; Hughes et al., 2017). This evolutionary process is influenced by multiple factors, including drug pressure and pathogen genetics. Firstly, the drug type and drug concentration can affect the nature and relative frequencies of AMR mutations observed in a given pathogen population (also known as the mutational profile for AMR) (Zhou et al., 2000; Ford et al., 2013; Lindsey et al., 2013; McGrath et al., 2014; Hughes et al., 2017; Huseby et al., 2017). Secondly, pathogen populations comprise genetically distinct strains, and this genetic variation may also influence AMR evolution (Fenner et al., 2012; Vogwill et al., 2014; Vogwill et al., 2016; Gagneux, 2018). Different pathogen genetic backgrounds can have different baseline susceptibilities to a given drug (Ängeby et al., 2010; Coeck et al., 2016), which can consequently affect patient treatment outcomes (Colangeli et al., 2018). The genetic background has also been shown to modulate the acquisition and prevalence of AMR (Borrell et al., 2009; Fenner et al., 2012; Ford et al.,

2013; Wollenberg et al., 2017), the mutational profile for AMR (Fenner et al., 2012; Ford et al., 2013; Vogwill et al., 2014; Oppong et al., 2019), and the phenotypic effects of AMR mutations (Gagneux et al., 2006c; Decuypere et al., 2012; Angst et al., 2013; Vogwill et al., 2016). Studying the interplay between pathogen genetics and drug pressure is therefore important in understanding how to restrict the emergence of AMR in pathogen populations.

AMR in *Mycobacterium tuberculosis* (*Mtb*), the etiological agent of human tuberculosis (TB), is of particular importance. *Mtb* infections globally cause the highest rate of mortality due to a single infectious agent both in general, and due to AMR specifically (WHO, 2018). Although the genetic variation in *Mtb* is small compared with other bacterial pathogens (Comas et al., 2010; Gagneux, 2018), several studies have shown that this limited genetic variation influences AMR phenotypes and prevalence (Gagneux et al., 2006c; Zaczek et al., 2009; Fenner et al., 2012; Gagneux, 2018). The global genetic diversity of *Mtb* comprises seven phylogenetic lineages (Comas et al., 2010; Gagneux, 2018), and *Mtb* strains belonging to the Lineage 2 Beijing/W genetic background have repeatedly been associated with multidrug-resistant TB (MDR-TB; defined as an infection from an *Mtb* strain that is resistant to at least isoniazid and rifampicin) both in vitro and in clinical settings (Borrell et al., 2009; Fenner et al., 2012; Ford et al., 2013; Merker et al., 2015; Wollenberg et al., 2017).

One strategy to reduce the emergence of AMR in *Mtb* is the development of new, shorter treatment regimens (Imperial et al., 2018; Vjecha et al., 2018). Many such experimental regimens use third- or fourth-generation fluoroquinolones (FQ) against drug-susceptible *Mtb* (Gillespie et al., 2014; Jindani et al., 2014; Merle et al., 2014; Imperial et al., 2018; Vjecha et al., 2018). However, FQs have long been integral to treating MDR-TB (Takiff et al., 2011), and the previous use of FQs has led to the emergence of FQ-resistance (FQ-R) in clinical strains of *Mtb* (Takiff et al., 1994; Maruri et al., 2012; Shah et al., 2017). FQ-R is one of the defining properties of extensively drug-resistant TB (XDR-TB), and XDR-TB accounts for 8.5% of MDR-TB cases (WHO, 2018). Understanding how FQ-R is acquired in natural populations of *Mtb* may allow for the development of tools or strategies to mitigate further increases in FQ-R prevalence.

In *Mtb*, the sole target of FQ is DNA gyrase (Takiff et al., 1994; Zhou et al., 2000; Piton et al., 2010; Aldred et al., 2016; Blower et al., 2016). Consequently, clinically relevant FQ-R in *Mtb* is primarily due to a limited set of chromosomal mutations located within the "quinolone-resistance-determining region" (QRDR) of the *gyrA* and *gyrB* genes, which encode DNA gyrase (Takiff et al., 1994; Maruri et al., 2012; Wollenberg et al., 2017). No horizontal gene-transfer or plasmid-based resistance to FQ has been documented in *Mtb* (Boritsch et al., 2016; Gygli et al., 2017). Studying FQ-R evolution in *Mtb* populations thus provides a promising setting for elucidating how the genetic background may affect the emergence and maintenance of clinically relevant chromosomal AMR mutations in bacterial populations.

While a great deal of literature exists on the biochemical mechanisms leading to the FQ-R phenotype in *Mtb* (Zhou et al., 2000; Piton et al., 2010; Mustaev et al., 2014; Aldred et al., 2016; Blower

et al., 2016), little is known on the evolutionary dynamics of FQ-R in different populations of *Mtb*. Given that antimicrobial regimens against *Mtb* infections use standardized, empirical dosing strategies (WHO, 2018), it is unclear whether different *Mtb* genetic backgrounds would acquire FQ-R at the same frequency when exposed to the same antimicrobial concentration. Whether the *Mtb* genetic background would also modulate the mutational profile for FQ-R, or the phenotypic effects of FQ-R mutations, is unknown. Such knowledge may provide insights on how to maintain or prolong the efficiency of FQs against different genetic variants of *Mtb* in the clinic.

In this study, we tested whether the *Mtb* genetic background plays a role in the evolution of FQ-R. We showed that the *Mtb* genetic background can lead to differences in the frequency of FQ-R emergence that span two orders of magnitude, as well as substantially modulate the mutational profile for FQ-R. We further demonstrated that the phenotypic effects of clinically relevant FQ-R mutations differed depending on the *Mtb* genetic background they were present in. Analysis of publicly available genomic sequences from clinical *Mtb* isolates also revealed a positive association between the FQ-R mutational profiles observed in vitro and the mutational profiles observed in the clinic. Taken together, we showed that the *Mtb* genetic background had a considerable role in evolution of FQ-R in the clinic.

3.3 Methods

3.3.1 Collection of Drug-Susceptible Clinical Isolates of *M. tuberculosis* Strains for In Vitro Studies

We used nine genetically distinct *Mtb* strains, with three strains from each of the following *Mtb* lineages: Lineage 1 (L1; also known as the East-Africa and India Lineage), Lineage 2 (L2; the East Asian Lineage), and Lineage 4 (L4; the Euro-American Lineage) (Comas et al., 2010; Gagneux, 2018). All strains were previously isolated from patients, fully drug-susceptible, and previously characterized by Borrell et al., 2019 (Table 7.1, see Supplementary Information). Prior to all experimentation, starter cultures for each *Mtb* strain were prepared by recovering a 20 μl aliquot from frozen stocks into a 10 ml volume of Middlebrook 7H9 broth (BD), supplemented with an albumin (Fraction V, Roche), dextrose (Sigma–Aldrich), catalase (Sigma–Aldrich), and 0.05% Tween 80 (AppliChem) (hereafter designated as 7H9 ADC). These starter cultures were incubated until their optical density at wavelength of 600 nm (OD_{600}) was $\sim$0.50, and were then used for in vitro assays.

3.3.2 Fluctuation Analyses

Fluctuation analyses were performed as described by Luria and Delbrück (Luria et al., 1943). Briefly, an aliquot from the starter cultures for each strain was used to inoculate 350 ml of 7H9 ADC to have an initial bacterial density of 5,000 colony forming units (CFU) per milliliter. This was immediately divided into 33 parallel cultures, each with 10 ml of culture volume aliquoted into individual 50 ml Falcon Conical Centrifuge Tubes (Corning Inc.). The parallel cultures were incubated at 37 °C on standing racks, with resuspension by vortexing (Bio Vortex V1, Biosan) every 24 h. Cultures were grown until

an OD_{600} of between 0.40 and 0.65. Once at this density, final cell counts (N_t) from three randomly chosen parallel cultures were calculated by serial dilution and plating on Middlebrook 7H11 (BD), supplemented with oleic acid (AppliChem), albumin, and catalase (hereafter referred to as 7H11 OADC). To calculate the number of resistant colonies (r), the remaining 30 parallel cultures not used for N_t determination were pelleted at 800 $\times$ g for 10 min at 4 °C using the Allegra X-15R Benchtop Centrifuge (Beckmann Coulter). The supernatants were discarded, and the bacterial pellets resuspended in 300 μl of 7H9 ADC. The resuspensions were spread on 7H11 OADC plates supplemented with the relevant drug concentration (2, 4, or 8 μg/ml of ofloxacin, or 100 μg/ml STR; Sigma). Resistant colonies were observed and enumerated after 21–35 days of incubation, depending on the *Mtb* strain. The frequency of drug-resistant mutants per culture (r) was enumerated, and the estimated number of drug-resistance mutations per culture (m) was estimated from the distribution of the r values (r_{dist}) using the Ma, Sarkar, Sandri-Maximum Likelihood Estimator method (MSS-MLE) (Rosche et al., 2000). Values of r that were >300 were simply given a value of 300, as this would not change the precision of the calculated m value using the MSS-MLE method (Rosche et al., 2000). The MSS-MLE method is also only valid for a range of m values between 0.3 and 20 (Rosche et al., 2000). The frequency of drug-resistance mutations acquired per cell (F) per strain was then calculated by dividing the calculated m values by their respective N_t values. The 95% confidence intervals for each F were calculated as previously described by Rosche and Foster (Rosche et al., 2000). Hypotheses testing for significant differences between the r_{dist} between strains for the fluctuation analyses at 4 μg/ml of OFX (Figure 3.1A) and at 100 μg/ml of STR (Figure 3.2) were performed using the Kruskal–Wallis test; significant differences in the r_{dist} between strains in the fluctuation analyses at 2 and 8 μg/ml of OFX (Figure 3.1B) were tested for using the Wilcoxon rank-sum test. Statistical analyses were performed using the R statistical software (v.3.5.1) (R Core Team, 2018).

3.3.3 Determining the Mutational Profile for Ofloxacin-Resistance In Vitro

From the parallel cultures plated on 4 μg/ml of OFX (Figure 3.1A), up to 120 resistant colonies per strain (at least 1 colony per plated parallel culture if colonies were present, to a maximum of 6) were transferred into 100 μl of sterile deionized H_2O placed in Falcon 96-well Clear Microplate (Corning Inc.). The bacterial suspensions were then heat-inactivated at 95 °C for 1 h, and used as PCR templates to amplify the QRDR in *gyrA* and *gyrB* using primers designed by Feuerriegel et al., 2009. PCR products were sent to Macrogen, Inc. or Microsynth AG for Sanger sequencing, and QRDR mutations were determined by aligning the PCR product sequences against the H37Rv reference sequence (Cole et al., 1998). Sequence alignments were performed using the Staden Package (Staden, 1996), while the amino acid substitutions identification were performed using the Molecular Evolutionary Genetics Analysis Version 6.0 package (Tamura et al., 2013). Fisher's exact test was used to test for significant differences between the strains' mutational profiles for OFX-R. Data analyses were performed using the R statistical software (v.3.5.1) (R Core Team, 2018), and figures were produced using the ggplot2

package (Wickham, 2016).

3.3.4 Isolation of Spontaneous Ofloxacin-Resistant Mutants

Spontaneous OFX-resistant mutants were isolated from strains belonging one of three genetic backgrounds: N0157 (L1, Manila sublineage), N1283 (L4, Ural sublineage), and N0145 (L2, Beijing sublineage) (Borrell et al., 2019). To begin, we transferred 50 μl of starter cultures for each strain into separate culture tubes containing 10 ml of fresh 7H9 ADC. Cultures were incubated at 37 °C until OD_{600} of $\sim$0.80, and pelleted at 800 $\times$ g for 5 min at 4 °C. The supernatant was discarded, and the pellet resuspended in 300 μl of 7H9 ADC. The resuspension was plated on 7H11 OADC (BD) supplemented with 2 μg/ml of OFX, and incubated until resistant colonies appeared ($\sim$14–21 days). Resistant colonies were picked and resuspended in fresh 10 ml 7H9 ADC, and incubated at 37 °C. Once the culture reached early stationary phase, two aliquots were prepared. The first aliquot was heat-inactivated at 95 °C for 1 h, and the *gyrA* mutation identified by PCR and Sanger sequencing, as described in the mutational profile for OFX-R assay. If the first aliquot harbored one of four OFX-R *gyrA* mutations ($GyrA^{D94G}$, $GyrA^{D94N}$, $GyrA^{A90V}$, $GyrA^{G88C}$), the second aliquot was stored in -80 °C for future use.

Prior to further experimentation with the spontaneously OFX-resistant mutant strains, starter cultures were prepared in the same manner as for the drug-susceptible strains.

3.3.5 Drug Susceptibility Assay

We determined the OFX-susceptibility levels of our spontaneous OFX-resistant mutants and their respective drug-susceptible ancestors by performing the colorimetric, microtiter plate-based Alamar Blue assay (Franzblau et al., 1998). Briefly, we used a Falcon 96-well Clear Microplate, featuring a serial 2-fold dilution of OFX. For drug-susceptible strains, a range of OFX concentration from 15 to 0.058 μg/ml was used. Meanwhile, for OFX-resistant strains, a range of 60–0.234 μg/ml was used. Each well was inoculated with a 10 μl volume of starter culture to have a final inoculum of $\sim 5 \times 10^6$ CFU/ml. The plates were incubated at 37 °C for 10 days. Following incubation, 10 μl of Resazurin (Sigma) were added to each well, and the plates were incubated for another 24 h at 37 °C. After this incubation period, plates were inactivated by adding 100 μl of 4% formaldehyde to every well. Measurement of fluorescence produced by viable cells was performed on SpectraMAX GeminiXPS Microplate Reader (Molecular Devices). The excitation wavelength was set at 536 nm, and the emission wavelength at 588 nm was measured. Minimum inhibitory concentration (MIC) for OFX was determined by first fitting a Hill curve to the distribution of fluorescence, and then defining the MIC as the lowest OFX concentration where the fitted Hill curve showed a $\geq$95% reduction in fluorescence. Two sets of experiments were performed for every strain, with three technical replicates per experiment. Analyses of MIC data were performed and figures created using the numpy, scipy, pandas, and matplotlib modules for the Python programming language.

3.3.6 Cell Growth Assay

We set up three or four 1,000 ml roller bottles with 90 ml of 7H9 ADC and 10 ml borosilicate beads. Each bottle was inoculated with a volume of starter cultures so that the initial bacterial density was at an OD_{600} of 5×10^{-7}. The inoculated bottles were then placed in a roller incubator set to 37 °C, and incubated for 12–18 days with continuous rolling. OD_{600} measurements were taken once or twice every 24 h. Two independent experiments in either triplicates or quadruplicates were performed per strain.

We defined the exponential phase as the bacterial growth phase where we observed a $\log_2$-linear relationship between OD_{600} and time; specifically, we used a Pearson's R^2 value $\geq$0.98 as the threshold. The growth rate of a particular strain was then defined as the slope of the linear regression model. The relative fitness of a given spontaneous OFX-R mutant was defined by taking the growth rate of the OFX-resistant mutant strain and dividing it by the growth rate of its respective drug-susceptible ancestor. Linear regression models for the cell growth assays data were performed using the numpy, scipy, pandas, and matplotlib modules for the Python programming language, as well as the R statistical software (v.3.5.1) and the ggplot2 package (Wickham, 2016; R Core Team, 2018).

3.3.7 Surveying the Fluoroquinolone-resistance Profile from Publicly Available *M. tuberculosis* Genomes

We screened public databases to download global representatives of *Mtb* genomes, as described by Menardo et al., 2018. We selected genomes that were classified as MDR-TB based on the presence of both INH-R and RIF-R mutations. This provided a data set of 3,450 genomes with confirmed MDR-TB. Due to excessive length, their accession numbers are only reported in the Supplementary Material of Castro et al., 2019 (number of genomes = 3,450). These MDR-TB genomes were then screened for the presence of FQ-R mutations, and we identified 854 genomes that were classified as FQ-R.

The INH-R, RIF-R, and FQ-R mutations used for screening are the same mutations used by Payne et al., 2019a. List of these mutations are only available in the Supplementary Material of Payne et al., 2019a and Castro et al., 2019, and not reported here due to excessive length (n = 196). A drug-resistance mutation was defined as "fixed" in the population when it reached a frequency of $\geq$90%. Meanwhile, a drug-resistance mutation was considered "variable" in the population when its frequency was between 10% and 90%; thus, multiple drug-resistance mutations may be present in the genomic data from a single *Mtb* clinical isolate.

3.3.8 Defining Transmission Clusters and Determining the Frequency of Fluoroquinolone-resistance *gyrA* Mutation Events

To define transmission clusters, the differences in the number of SNPs were used as a measure of genetic distance between two *Mtb* genomes. Using the haplotypes package (v.1.0) for the R statistical software (v.3.5.1) (R Core Team, 2018), a genetic distance matrix was then inferred for the 3,450 MDR-TB genomes. Insertions/deletions were considered missing data. Agglomerative clustering was performed using the agnes function from the cluster package (v.2.0.6) for the R statistical software (R Core Team, 2018). A conservative threshold of 12 SNPs average distance was used to define likely patient-to-patient transmission (Walker et al., 2013), and the tree was cut at a height of 12 SNPs using the hclust function. All resulting transmission clusters, with a minimum size of two clustered genomes, were used for further analysis. For every transmission cluster, each unique and fixed FQ-R *gyrA* mutation was treated as an independent mutation event. Fixed FQ-R *gyrA* mutations in nonclustered *Mtb* genomes were also treated as an independent mutation event. Figures were produced using the ggplot2 package (Wickham, 2016).

3.4 Results

3.4.1 Frequency of Ofloxacin-Resistance in *M. tuberculosis* Is Strain-Dependent

We first tested for whether the *Mtb* genetic background led to differences in the frequency of FQ-R acquisition. To do so, we performed a Luria–Delbrück fluctuation analysis on nine drug-susceptible and genetically distinct *Mtb* clinical strains belonging to Lineage 1 (L1), Lineage 2 (L2), and Lineage 4 (L4) (Table 7.1, see Supplementary Information) (Luria et al., 1943; Comas et al., 2010; Gagneux, 2018; Borrell et al., 2019). We measured their frequency of resistance in vitro to the FQ ofloxacin (OFX), as OFX was used extensively to treat MDR-TB patients in the past (Takiff et al., 2011). Given that anti-TB treatment regimens use standardized drug concentrations (WHO, 2018), we also measured the frequency of resistance to the same concentration of OFX (4 μg/ml) for all nine strains. We observed significant strain-dependent variation in the frequency of OFX-resistance (OFX-R) at 4 μg/ml, with the difference spanning two orders of magnitude (Figure 3.1A, $P = 2.2 \times 10^{-16}$, Kruskal–Wallis). Several of the nine drug-susceptible *Mtb* strains contained missense substitutions in DNA gyrase that are not associated with FQ-R (Table 7.2, see Supplementary Information) (Borrell et al., 2019). These mutations are phylogenetic markers that reflect the population structure of *Mtb* and cannot be avoided if strains from different *Mtb* lineages are used (Comas et al., 2010; Gagneux, 2018). We found no evidence for any associations between the presence a given phylogenetic DNA gyrase missense mutation and the frequency of OFX-R acquired.

The concentration of the antimicrobial can affect the observed frequencies of AMR in *Mtb* (Zhou et al., 2000; Ford et al., 2013; McGrath et al., 2014). Therefore, we tested whether changing the selective concentration of OFX would affect the relative differences in strain-specific OFX-R frequencies. For

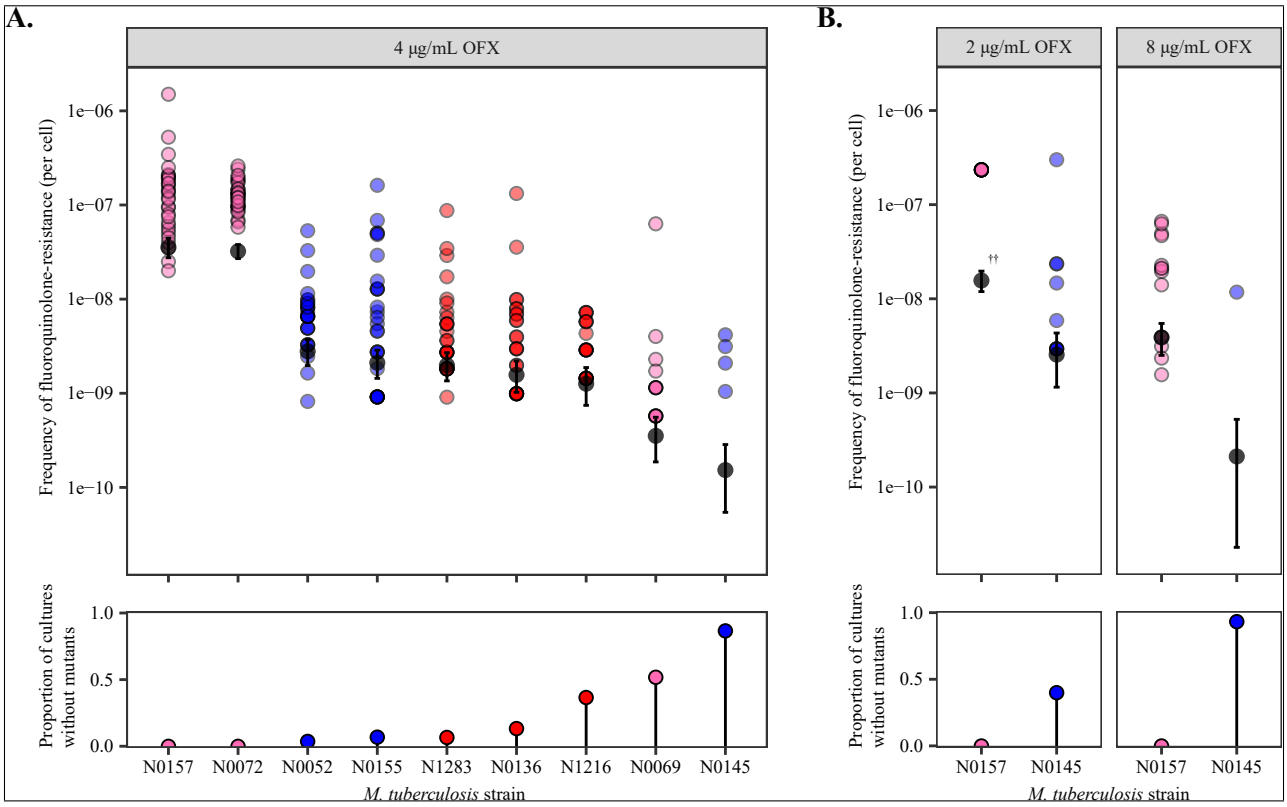


Figure 3.1: Variation in the frequency of ofloxacin-resistance between genetically distinct, wild-type *Mycobacterium tuberculosis* strains.
(A) Frequency of ofloxacin-resistance at 4 μg/ml ofloxacin (OFX), as measured by fluctuation analysis. Top panel: colored points represent the frequency of ofloxacin-resistance per cell per parallel culture, with darker points representing multiple cultures with the same frequency. Colors denote the lineage that the *Mtb* strain belongs to (L1, pink; L2, blue; L4, red). Gray points represent the estimated number of drug-resistance mutations per cell per strain as calculated by MSS-MLE, while black bars denote the respective 95% confidence intervals. Bottom panel: the percentage of parallel cultures lacking OFX-resistant mutants. (B) Frequency of ofloxacin-resistance at 2 or 8 μg/ml OFX. †† Due to restrictions on the range of values that the MSS-MLE method is valid for (see Methods), the estimated number of drug-resistance mutations per cell for N0157 at 2 μg/ml OFX, as presented here, is an underestimate.

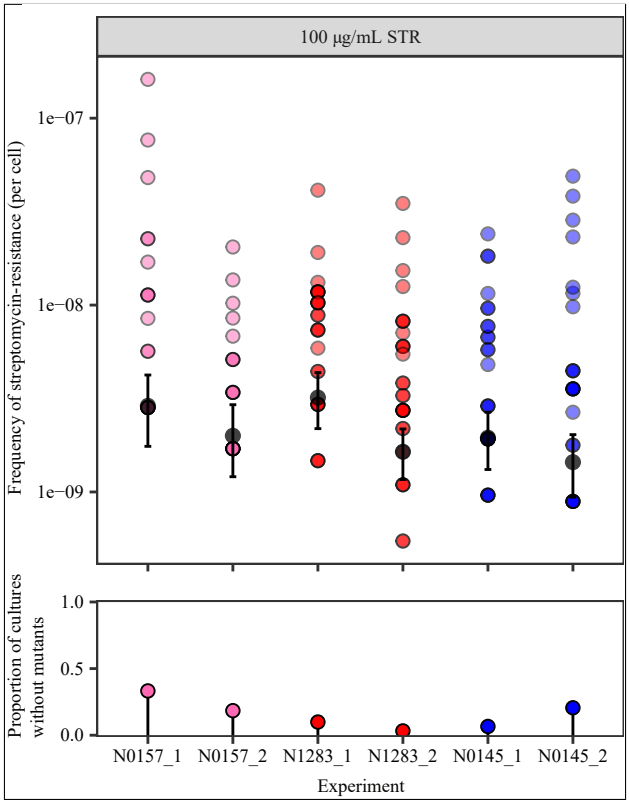


Figure 3.2: The frequency of streptomycin-resistance at 100 μg/ml streptomycin (STR) for wild-type N0157, N1283, and N0145 *Mycobacterium tuberculosis* strains, as measured by fluctuation analysis. Top panel: colored points represent the frequency of streptomycin-resistance per cell per parallel culture, with darker points representing multiple cultures with the same frequency. Colors denote the lineage that the *Mtb* strain belongs to (L1, pink; L2, blue; L4, red). Gray points represent the estimated number of drug-resistance mutations per cell per strain as calculated by MSS-MLE, while black bars denote the respective 95% confidence intervals. Bottom panel: the percentage of parallel cultures lacking STR-resistant mutants. Two biological replicates are presented for each *Mtb* strain, with each replicate identifier suffixed after the strain name.

the sake of simplicity, we tested only two strains, with each strain at the opposite extremes of the frequency of resistance to 4 μg/ml OFX, as shown in Figure 3.1A: N0157 (high OFX-R frequency) and N0145 (low OFX-R frequency). We found that the frequency of OFX-R remained one to two orders of magnitude higher in N0157 than in N0145 across all the concentrations we tested (Figure 3.1B, $P = 2.46 \times 10^{-5}$ for 2 μg/ml OFX, and $P = 4.03 \times 10^{-6}$ for 8 μg/ml OFX, Wilcoxon rank-sum test). The N0157 strain had nearly confluent growth at 2 μg/ml OFX, which is the OFX concentration that has been shown to inhibit 95% of *Mtb* strains that have not been previously exposed to OFX, but does not inhibit *Mtb* strains that are considered resistant to OFX in the clinic (Ängeby et al., 2010; Coeck et al., 2016). This suggested that N0157 has low-level resistance to OFX, despite having no mutation in the QRDR. Meanwhile, at 8 μg/ml OFX, we observed only four resistant colonies for N0145 across all samples, with all colonies arising within the same culture.

The variation in OFX-R frequencies when selecting on the same concentration of OFX may be

driven by several, non-exclusive biological factors. Firstly, the *Mtb* strains we tested may have different baseline DNA mutation rates. Secondly, the number and relative frequency of potential mutations that confer OFX-R may vary depending on the *Mtb* genetic background. Thirdly, the relative cost of OFX-R mutations may differ between *Mtb* genetic backgrounds. As the observed frequency of OFX-R in *Mtb* is likely the result from a combination of multiple factors, we took advantage of the fact that we had identified strains with a range of OFX-R frequencies. We selected three representative strains with significantly different OFX-R frequencies: N0157, N1283, and N0145. These strains had a high-, mid-, and low frequency of OFX-R, respectively (Figure 3.1A). We then explored the relative contributions of each biological factor listed above in driving the variation in OFXR across genetically distinct *Mtb* strains.

3.4.2 Mutation Rate Differences Do Not Drive the In Vitro Variation in Ofloxacin-Resistance Frequency in *M. tuberculosis*

We first tested for the presence of differential mutation rates between our panel of *Mtb* strains in Figure 3.1A. Mutations in *dnaE*, which encodes the replicative DNA polymerase and serves as the major replicative exonuclease in *Mtb*, have been shown to confer a hypermutator phenotype in *Mtb* in the absence of environmental stress (Rock et al., 2015; Baños-Mateos et al., 2017). While *dnaE* mutations were present in the genomic data of our panel of drug-susceptible *Mtb* strains (Table 7.2, see Supplementary Information) (Borrell et al., 2019), none was in the polymerase and histidinol phosphatase domain of DnaE, the region where mutations would impart a hypermutator phenotype (Rock et al., 2015; Baños-Mateos et al., 2017). We did not test for the presence of *dnaE* mutations in the resistant colonies following the fluctuation analysis, as we reasoned that the likelihood of gaining both a *dnaE* and a *gyrA* double mutation within this relatively short period is extremely low as to be considered negligible. To test for mutation rate variation in vitro, we again conducted a fluctuation analysis on N0157, N1283, and N0145 (the high-, mid-, and low-frequency OFX-R strains, respectively), but used streptomycin (STR; 100 μg/ml) instead of OFX. We hypothesized that if the frequency of OFX-R is driven by differential mutation rates, then we should expect similar differences in the frequency of STR-resistance (STR-R). However, we observed no evidence for differences in the frequency of STR-R between the strains tested (Figure 3.2, $P = 0.135$, Kruskal–Wallis; Table 7.3, see Supplementary Information). This suggested that the observed differences in frequency of resistance are specific to OFX, and that there are limited, if any, inherent differences in mutation rates between the *Mtb* strains tested.

3.4.3 Mutational Profile for Ofloxacin-Resistance Is Highly Strain-Dependent

We next determined the mutational profile for OFX-R for each strain used in the fluctuation analysis at 4 μg/ml OFX (Figure 3.1A). The QRDR mutations in 680 *gyrA* and 590 *gyrB* sequences were identified in the resistant colonies. We observed that *gyrA* mutations made up 99.7% of the QRDR

mutations observed (645 *gyrA* mutations, 2 *gyrB* mutations), and no QRDR double mutants were present (Figure 7.1, Tables 7.4 & 7.5, see Supplementary Information). The mutational profiles for OFX-R were also highly strain-specific (Figure 3.3A, $P = 5.00 \times 10^{-4}$, Fisher's exact test). Specifically, the GyrA A90V mutation was most prevalent in the high-frequency OFX-R strains, while GyrA D94G was most prevalent in all other strains. There was also a slight trend showing that strains with a greater number of unique *gyrA* mutations present also had higher rates of OFX-R (Figures 3.1A & 3.3B).

The strain-dependent variation in the mutational profile for OFX-R may be due to *gyrA* mutations conferring different resistance levels depending on the *Mtb* strain they are present in. To test this hypothesis, we first isolated OFX-R mutants carrying one of four possible GyrA mutations (G88C, A90V, D94G, or D94N) in the three strains used in Figure 3.2: N0157, N1283, and N0145. The OFX MIC was determined for each of the 12 OFX-R mutant strains, along with their respective wild-type ancestors. We found that each parental wild-type strain had different susceptibilities to OFX, with N0157, N1283, and N0145 having OFX MICs of 2, 0.6, and 0.5 μg/ml, respectively (Figure 3.4A and Table 7.6, see Supplementary Information). This was consistent with the fluctuation analysis results shown in (Figure 3.1B). Furthermore, we observed that the OFX MIC conferred by a given *gyrA* mutation varied depending on the strain it was present in (Figure 3.4B and Table 7.6, see Supplementary Information). For example, mutants in the N0157 strain generally had higher OFX MICs than mutants in either the N0145 or N1283 strains. The only mutation that deviated from this trend was GyrA G88C, which conferred a higher OFX MIC when in the N0145 strain. Notably, the GyrA A90V mutation conferred a resistance level $\geq$4 μg/ml OFX in the N0157 and N1283 strains, but not in N0145. This was consistent with the presence of GyrA A90V in the OFXR mutational profile for N0157 and N1283, but not in N0145, in the fluctuation analysis using 4 μg/ml OFX (Figures 3.1A and 3.3). In summary, the differences in OFX MIC reflected the strain-dependent mutational profiles for OFX-R in *Mtb*, as expected.

3.4.4 Fitness of Ofloxacin-Resistance Mutations Are Associated with Their Relative Frequency In Vitro

While the OFX MICs may determine which mutations may be observed in a fluctuation analysis, it is not the sole parameter to influence the OFX-R mutational profile for a given strain. We found that while the same *gyrA* mutation can be observed in two different *Mtb* strains, their relative frequencies may vary (Figure 3.3). This variation may be due to the fitness of a given *gyrA* mutant being different across genetic backgrounds. To test this hypothesis, we used cell growth assays in antibiotic-free conditions to measure the in vitro fitness of our panel of OFX-R mutants relative to their respective parental wild-type ancestors. We observed that the relative fitness of the OFX-R mutants was modulated by both the *gyrA* mutation and the *Mtb* strain they were present in (Figure 3.5A; Figures 7.2 & 7.3, Table 7.7, see Supplementary Information). Furthermore, there was a positive association between the

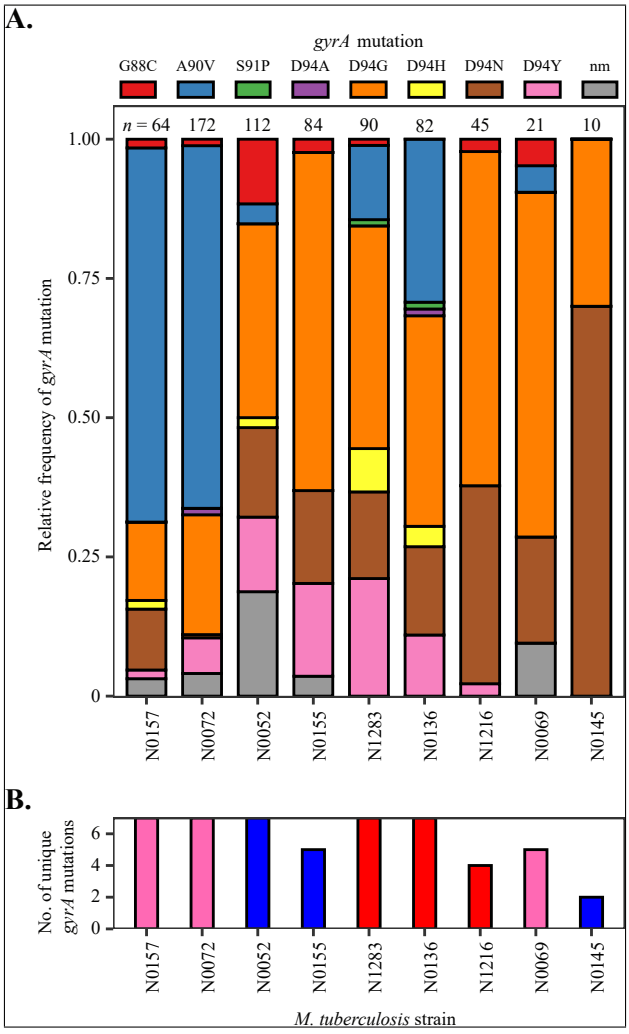


Figure 3.3: Variation in the mutational profile for ofloxacin-resistance after fluctuation analyses using nine genetically distinct *Mycobacterium tuberculosis* strains.
(A) Mutations in the quinolone-resistance-determining region (QRDR) of *gyrA* were analyzed in 680 ofloxacin (OFX)-resistant colonies from the fluctuation analysis performed in Figure 3.1A (nm, no identified QRDR *gyrA* mutations). Strains are ordered left to right based on their frequency of OFX-resistance at 4 μg/ml OFX. Numbers of colonies analyzed per strain are reported directly above each column. (B) The number of unique QRDR *gyrA* mutations per *Mtb* strain for OFX-resistance. Bar colors denote the *Mtb* lineage the strain belongs to (L1, pink; L2, blue; L4, red).

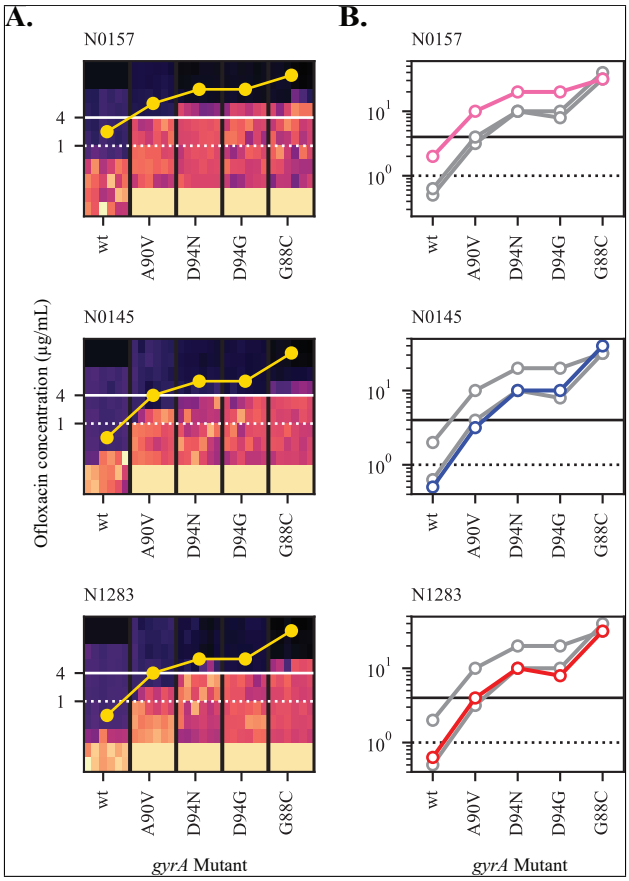


Figure 3.4: The *Mycobacterium tuberculosis* genetic background modulates the ofloxacin (OFX) minimum inhibitory concentration (MIC).
(A) Heat-map of OFX-susceptibility via Alamar Blue assay for *gyrA* mutant strains of *Mtb*, as well as their wild-type ancestor, in three genetic backgrounds (N0157, N0145, or N1283). Light areas represent growing cultures, while dark areas represent nongrowing cultures. Yellow points represent estimates for OFX MIC (≥95% reduction in fluorescence). Areas of solid black colors (at 16+ μg/ml OFX for wild-type) and solid light beige colors (at <0.125 μg/ml OFX for mutants) were not measured and colored in for illustrative purposes. (B) OFX MIC estimates for each strain per genetic background, superimposed. Colored points and lines represent MIC measurements for highlighted genetic background, with the line color denoting the lineage that the strain belongs to (L1, pink, L2, blue, L4, red). Gray points and lines represent the other two genetic backgrounds. Solid horizontal black line denotes 4 μg/ml OFX, while dashed horizontal black line denotes 1 μg/ml OFX.

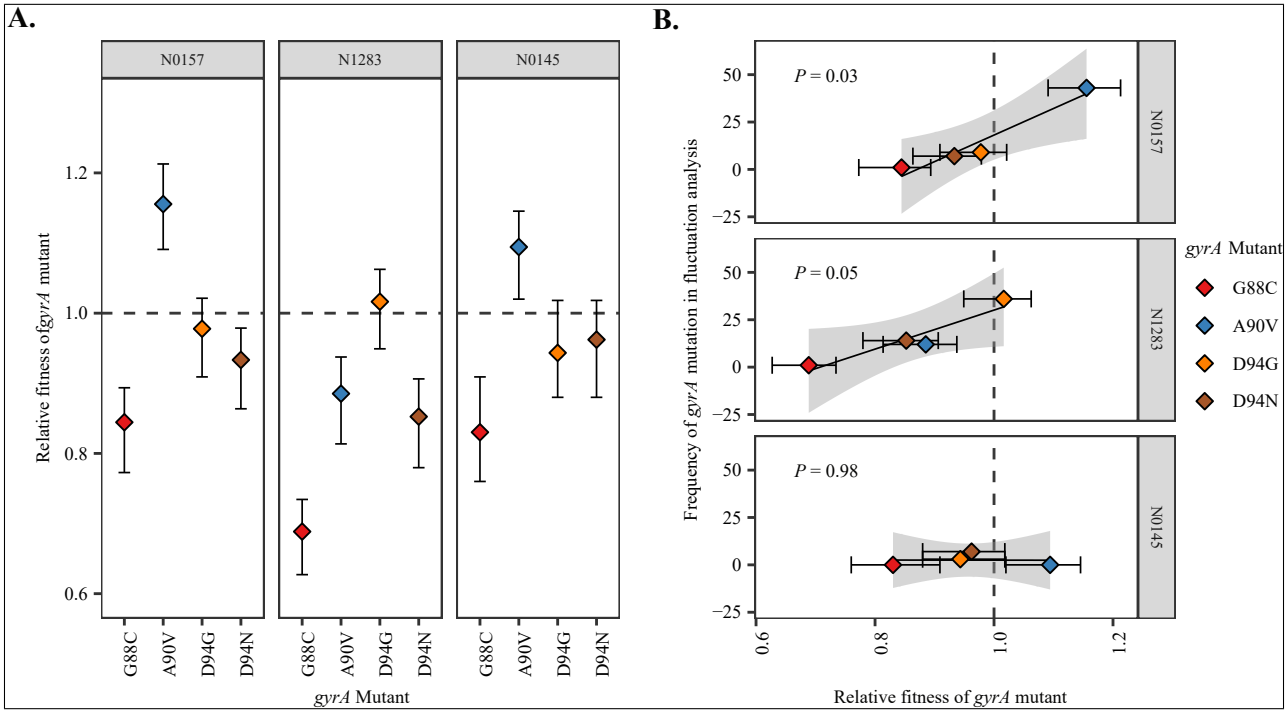


Figure 3.5: The *Mycobacterium tuberculosis* genetic background modulates the fitness effect of fluoroquinolone-resistance mutations.
(A) Fitness of ofloxacin-resistant *Mtb* strain with specified *gyrA* mutation relative to the fitness of their respective wild-type ancestral strain. Fitness was measured by cell growth assay in antibiotic-free conditions. Ancestral strain per *gyrA* mutant is indicated in the gray bar above each panel. (B) Association between the relative fitness of specified *gyrA* mutant and their absolute frequency after the fluctuation analysis performed in Figure 3.1A and as reported in Figure 3.3A, in three genetic backgrounds (N0157, N1283, and N0145).

fitness of a given *gyrA* mutation with its relative frequency in the fluctuation analysis for the N0157 and N1283 strains (Figure 3.5B, $P = 0.03$ for N0157, $P = 0.05$ for N1283). There was no evidence of an association in the N0145 background due to the lack of GyrA G88C and A90V mutants in its fluctuation analysis.

The results from Figures 3.4 and 3.5, as well as the apparent lack of mutation rate differences between our strains (Figure 3.2), suggested that differential mutational profiles were an important contributor in the variation in OFX-R frequency in *Mtb*. These mutational profile differences appear to be driven by the *Mtb* genetic background's effect on both the MIC and the relative fitness cost of OFX-R mutations. We next explored whether these in vitro results would be relevant in clinical settings.

3.4.5 Mutational Profile for Fluoroquinolone-resistance In Vitro Reflects Clinical Observations

To explore the clinical relevance of our in vitro work, we surveyed the FQ-R mutational profile from publicly available *Mtb* genomes obtained from clinical isolates. FQs are generally used for treat-

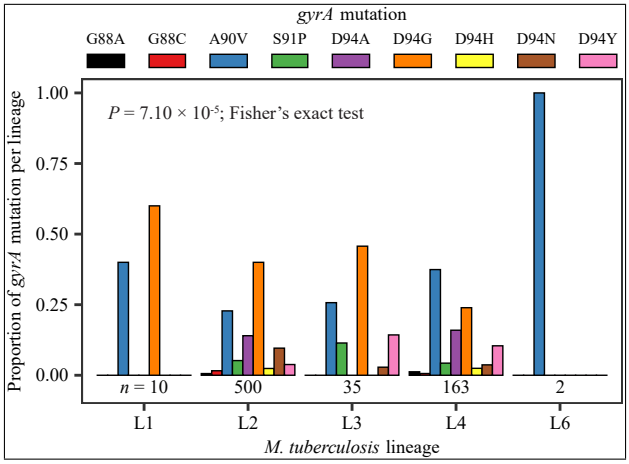


Figure 3.6: Mutational profile for fluoroquinolone-resistance *gyrA* mutations is lineage-specific in clinical isolates of *Mycobacterium tuberculosis*.
An initial data set consisting of 3,450 genomes with confirmed MDR-TB mutations were surveyed. About 854 genomes were identified as fluoroquinolone-resistant, with 848 of these genomes containing at least one *gyrA* mutation. Only fixed fluoroquinolone-resistance mutations in the *gyrA* gene are enumerated here (n = 710). No fixed mutations were observed in Lineage 5 strains. Numbers of genomes analyzed per lineage are presented directly below their respective bar graph.

ment against MDR-TB (WHO, 2018). While it is unclear whether resistance mutations for isoniazid (INH) and/or rifampicin (RIF) predispose a strain to become FQ-R, the prevalence of FQ-R is heavily biased toward MDR-TB strains due to treatment practices. We therefore based our analyses on a collated data set of 3,450 publicly available MDR-TB genomes (Castro et al., 2019), which we confirmed to be MDR-TB based on the presence of known INH-resistance (INH-R) and RIF-resistance (RIF-R) mutations. This data set provided a reasonable sampling of the overall genetic diversity of *Mtb*, as six of the seven known phylogenetic *Mtb* lineages were represented (Lineages 1–6) (Comas et al., 2010; Gagneux, 2018). We catalogued their FQ-R mutational profiles, and found 950 FQ-R mutations in 854 genomes (Tables 7.8 & 7.9, see Supplementary Information), showing that multiple FQ-R mutations may be present in the genome of a single *Mtb* clinical isolate. The frequency of FQ-R differed between lineages, with the highest frequencies present in L2 and L4 strains (Table 7.8, see Supplementary Information; $P < 2.20 \times 10^{-16}$, Chi-square Goodness of Fit Test). Moreover, we noticed a lineage-dependent mutational profile for FQ-R (Figure 3.6, $P = 7.10 \times 10^{-5}$, Fisher's exact test; Tables 7.9 & 7.10). For example, while the GyrA D94G mutation was most prevalent in strains belonging to L1, L2, and Lineage 3 (L3), the GyrA A90V mutation was most prevalent in L4 and Lineage 6 (L6).

We observed that the mutational profile for FQ-R in the fluctuation analysis experiments mimicked published clinical data. Firstly, *gyrA* mutations made up the large majority of FQ-R mutations in vitro (Figure 3.3; Tables 7.4 & 7.5, see Supplementary Information) and 944 out of the 950 QRDR mutations in the clinic (99.6%; Table 7.9 , see Supplementary Information). We then tested whether the *Mtb* genetic background had an impact on FQ-R *gyrA* mutational profiles in the clinic as it did in

vitro. However, transmission events can modulate the frequency of FQ-R *gyrA* mutations in the clinic, but not in a fluctuation analysis. If each genome from the clinical data is treated as an independent event, then number of FQ-R *gyrA* mutation events in the clinic would be overestimated compared with the number of mutation events in a fluctuation analysis. Therefore, rather than directly compare the frequency of *gyrA* mutations from the OFX-R mutational profile in Figure 3.3 to the absolute frequency of *gyrA* mutations in the genomic data survey, we instead compared the in vitro frequency of *gyrA* mutations in Figure 3.3 to the frequency of mutation events per *gyrA* mutation in the genomic data. To do so, we first used differences in the number of single-nucleotide polymorphisms (SNPs) as a measure of genetic distance between two genomes, then defined transmission clusters within the 3,450 MDR-TB genomes via a conservative cutoff of 12 SNPs average distance (Walker et al., 2013). Each unique and fixed FQ-R *gyrA* mutation present per transmission cluster, as well as each fixed FQ-R *gyrA* mutations present in nonclustered genomes, were counted as independent mutation events. We limited our analysis to L2 and L4 strains, as these two lineages had the highest clinical frequencies of FQR. We observed that the profile for FQ-R *gyrA* mutation events in L2 strains differed significantly to L4 strains (Figure 7.5, Table 7.11, see Supplementary Information; $P = 0.02$, Fisher's exact test). Furthermore, there was a positive association between the frequency of a given FQ-R *gyrA* mutation in our fluctuation analysis compared with the frequency of its mutation event in the clinic for both L2 and L4 strains (Figure 3.7, $P = 0.02$ for L2, $P = 0.04$ for L4). Based on the adjusted R^2 values, 51% of the variability in the clinical frequencies of FQ-R *gyrA* mutation events in L2 strains and 39% of the variability in L4 strains can be attributed to how FQ-R evolves in *Mtb* in vitro. As the in vitro evolution of FQ-R is itself modulated by the *Mtb* genetic background, this provided evidence for the *Mtb* genetic background's role in the evolution of FQ-R in the clinic.

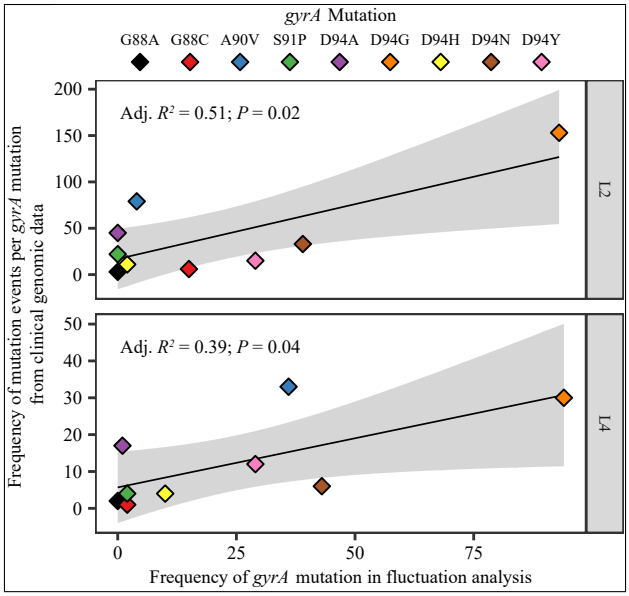


Figure 3.7: Association between the clinical frequency of mutation events of each fluoroquinolone-resistance (FQ-R) *gyrA* mutations with their respective in vitro frequencies among *Mycobacterium tuberculosis* strains belonging to either the L2 or L4 lineages.
Mutation events per FQ-R *gyrA* mutation were enumerated from an initial data set of 3,450 genomes with confirmed MDR-TB mutations. Each unique and fixed FQ-R *gyrA* mutation present per transmission cluster (cutoff = 12 SNPs average distance), as well as each fixed FQ-R *gyrA* mutation present in nonclustered genomes, were counted as independent mutation events. The in vitro frequencies of FQ-R *gyrA* mutations presented here are the same as in Figure 3.3A, grouped by lineage.

3.5 Discussion

Overall, we illustrate the *Mtb* genetic background's considerable role in the evolution of resistance to FQs, a clinically important antimicrobial. We first explored whether the genetic variation among natural populations of *Mtb* can influence FQ-R evolution in vitro. Specifically, considering that *Mtb* treatment regimens are based on standardized antimicrobial concentrations (WHO, 2018), we tested whether different genetic variants of *Mtb* would acquire FQ-R at the same frequency when exposed to the same concentration of FQ. Fluctuation analysis on nine, genetically distinct, drug-susceptible *Mtb* strains showed that the genetic background can have a drastic effect on the rate of OFX-R acquisition when using the same concentration of OFX (Figure 3.1). Our results provide the first evidence showing that the *Mtb* genetic background can modulate the frequency of FQ-R acquisition.

However, the effect of the *Mtb* genetic background on AMR frequencies observed here in the context of OFX-R differed from those reported in previous work focusing on other antibiotics. Past literature has focused on the positive association between MDR-TB and L2 Beijing (Borrell et al., 2009; Fenner et al., 2012; Merker et al., 2015; Eldholm et al., 2016). Initial genetic analysis on a global collection of strains showed that mutations in DNA repair genes were associated with being MDR-TB, and that these mutations were specific to L2 Beijing isolates (Rad et al., 2003). The authors thus hypothesized that L2 Beijing strains may have a hypermutator phenotype, which would lead to higher rates of AMR mutations acquisition (Rad et al., 2003). Based on this L2 Beijing hypermutator hypothesis, one would expect that L2 Beijing strains would also show higher frequencies of FQ-R. However, this was not the case in our fluctuation analysis for OFX-R, as one of our L2 Beijing strains (N0145) repeatedly acquired the lowest frequency of OFX-R (Figure 3.1). Moreover, we saw minimal, if any, DNA base-pair mutation rate differences between three *Mtb* strains (one of which was L2 Beijing) with different in vitro OFX-R frequencies (Figure 3.2). Our results therefore contradict the L2 Beijing hypermutator hypothesis. Published experimental work have also provided varying results. Initial fluctuation analyses showed no difference in the frequency of RIF-R in L2 Beijing strains compared with non-L2 Beijing strains (Werngren et al., 2003). In contrast, a fluctuation analysis performed by Ford et al. (2013) showed that L2 Beijing strains had higher frequencies of resistance for INH, RIF, and ethambutol compared with L4 strains, even after correcting for differences in AMR mutational profiles. Lastly, a more recent fluctuation analysis using the same concentrations of INH or RIF as Ford et al. showed that while different *Mtb* strains had different frequencies of INH-R, a L2 Beijing strain did not have higher frequencies of INH-R nor RIF-R compared with non-L2 Beijing strains (Carey et al., 2018). Although diverging in their results, these in vitro and genetic studies, together with the study conducted here, highlight the importance of the genetic background when testing for the frequency of AMR in *Mtb*. Furthermore, these results show that differential DNA mutation rates is not the only parameter relevant in determining the frequency of FQ-R in *Mtb*.

If DNA mutation rates do not contribute to the variation in OFX-R frequency, we hypothesized that differences in the phenotypic effects of OFX-R mutations, and their consequent effect on the

mutational profiles for OFX-R, may be important contributors. By sequencing the QRDR from resistant colonies in our OFX fluctuation analysis, we observed strain-specific patterns in the mutational profiles for OFX-R (Figure 3.3). This suggested that the mutational profile for FQ-R is not only a function of the FQ type and concentration (Zhou et al., 2000; Malik et al., 2010; Huseby et al., 2017) but that epistatic interactions between a given FQ-R mutation and the genetic background may also play a role. Similar epistasis between the phenotype of a given AMR mutation and the genetic background have been observed in other bacteria. For example, a given RIF-R *rpoB* mutation can confer differential MIC and fitness costs depending on the genetic background it occurred in, or on the presence of other AMR mutations, in *Escherichia coli* (Angst et al., 2013), *Pseudomonas* spp. (Vogwill et al., 2014; Vogwill et al., 2016), *Mycobacterium smegmatis* (Borrell et al., 2013), and *Mtb* (Gagneux et al., 2006c; Zaczek et al., 2009). In line with these previous studies, we found that the OFX MIC and the fitness effect conferred by a given *gyrA* mutation varied significantly depending on the *Mtb* genetic background they occur in (Figures 3.4 and 3.5A; Tables 7.6 & 7.7, see Supplementary Information). These results suggest that epistasis plays a role in determining the strain-dependent OFX-R frequencies and mutational profiles observed during our fluctuation analyses (Figures 3.3 and 3.5B).

The epistasis between the *Mtb* genetic background and FQ-R mutations may have clinical consequences. A recent study has shown that drug-susceptible *Mtb* strains with higher MICs to INH and RIF were associated with increased risk of relapse following first-line treatment (Colangeli et al., 2018). FQ-R *gyrA* mutations that confer higher MICs, such as any *gyrA* mutation in codon D94 except for D94A, have also been associated with poorer treatment outcomes in MDR-TB patients (Rigouts et al., 2016; Farhat et al., 2017). Considering our observation that the *Mtb* genetic background affected both the OFX MICs and OFX-R mutational profiles (Figures 3.3 and 3.4; Tables 7.4, 7.5, & 7.6, see Supplementary Information), the genetic background may therefore contribute to differences in patient treatment outcomes when using FQs as first-line drugs.

Using publicly available genomic data from *Mtb* clinical isolates, we observed significant lineage-dependent variation in the frequency of and mutational profiles for FQ-R (Figure 3.6). As expected, the vast majority of FQ-R mutations were observed in *gyrA* (Tables 7.9 & 7.10, see Supplementary Information) (Takiff et al., 1994; Zhou et al., 2000; Piton et al., 2010; Maruri et al., 2012; Aldred et al., 2016; Blower et al., 2016; Wollenberg et al., 2017). FQ-R was also most frequent in L2 and L4. This was also as expected, as strains from the L2 Beijing sublineage are known to associate with MDR-TB (Borrell et al., 2009; Fenner et al., 2012; Merker et al., 2015; Wollenberg et al., 2017), while L4 strains are the most prevalent globally, including in regions classified as high burden for TB (Stucki et al., 2016; WHO, 2018; Brynildsrud et al., 2018; Gagneux, 2018). Consequently, strains from L2 and L4 would be more exposed to FQs, leading to the higher FQ-R frequencies observed in these two lineages. Furthermore, we observed that more than half of the variability in the clinical frequency of FQ-R *gyrA* mutation events in L2 strains can be explained by how *Mtb* evolves in vitro (Figure 3.7). However, the in vitro FQ-R evolution could only account for 39% of the variability for the frequency of FQ-R

gyrA mutation events in clinical L4 strains. This suggested that while the *Mtb* genetic background can influence the evolution of FQ-R in the clinic, other factors (which may be independent of the *Mtb* genetic background) likely played strong roles as well. Epidemiological factors including socioeconomic disruptions, health system inefficiencies, and human behavior are well known risk factors for the emergence and transmission of AMR in *Mtb* (Dalton et al., 2012; Merker et al., 2015; Alvarez-Uria et al., 2016; Eldholm et al., 2016; Shah et al., 2017). Meanwhile, biological factors not explored in this study, such as antibiotic type and concentration (Zhou et al., 2000; Ford et al., 2013; Lindsey et al., 2013; McGrath et al., 2014; Mustaev et al., 2014; Malik et al., 2016), pharmacodynamic and pharmacokinetic features (Pienaar et al., 2017; Sarathy et al., 2018), and the selective pressure of the host immune system (Handel et al., 2009), may also influence the evolution of FQ-R.

How the genetic background modulates FQ-R evolution in the clinic may also differ between different bacterial pathogens. In contrast to *Mtb* where DNA gyrase is the sole target for FQs, Gram-negative bacteria such as *E. coli* and *Salmonella* have two targets for FQs: DNA gyrase and topoisomerase IV (Hooper, 1999). In *E. coli*, the evolutionary trajectory toward high-level FQ-R generally involves the stepwise acquisition of FQ-R mutations in either DNA gyrase or topoisomerase IV (Huseby et al., 2017). Therefore, the genetic background of Gram-negative bacteria may modulate the phenotypes of FQ-R mutations not only in DNA gyrase but in topoisomerase IV as well. Nevertheless, a recent study has shown that a common FQ-R mutation in Gram-negative bacteria, GyrA S83L, confers different phenotypes depending on whether it is present in *E. coli* or in *Salmonella* (Apjok et al., 2019). This suggests that the genetic background of Gram-negative bacteria may affect the evolution of FQ-R in the clinic. This type of epistasis is also not restricted to bacteria. The genetic background modulated the phenotypes of AMR mutations in the protozoan parasite *Leishmania donovani* (Decuypere et al., 2012), and the phenotypes of mutations in yeast when exposed to different environments, including antimicrobial exposure (Mullis et al., 2018). Thus, while the mode of epistasis between the genetic background and the phenotypes of mutations may differ in different organisms and environments, published work and the results of our study provide compelling evidence that this epistasis is a major factor in the evolution of AMR in both prokaryotic and eukaryotic organisms.

Our study is limited by the fact that our survey of clinical FQ-R frequencies involved a genomic data set that was sampled by convenience. This data set was used due to its public availability, and may not be fully representative of FQ-R frequencies in *Mtb* populations. We noted that lineage-specific frequencies of FQ-R were likely biased due to the overrepresentation L2 and L4 strains. Thus, to acquire a better understanding on which FQ-R mutations appeared and at what frequency they occurred at in different *Mtb* lineages, either more genomes from clinical isolates from other *Mtb* lineages must be made available, or a population-based study must be undertaken, preferably in a high-burden MDR-TB region.

Exposure to quinolones have been shown to lead to SOS response-mediated mutagenesis, which can increase the rate of AMR acquisition, including resistance to quinolones themselves (Cirz et al.,

2005; Malik et al., 2010; Frenoy et al., 2018). Therefore, the strain-dependent OFX-R acquisition rates Figure 3.1 may be due to strain-dependent differences in the magnitude of quinolone-induced mutagenesis. We did not explicitly test for this possibility. However, phylogenetic SNPs present in SOS response-related genes may lead to strain-dependent differences in quinolone-induced mutagenesis, and we observed no such SNPs present across our panel of drug-susceptible *Mtb* strains (Table 7.2, see Supplementary Information) (Borrell et al., 2019). Thus, we observed no genetic evidence for strain specific SOS response-mediated mutagenesis. Furthermore, in *E. coli*, quinolone-induced quinolone-resistant mutations may only be observed after 5 days of incubation with quinolones, which is equivalent to >220 generations for wild-type *E. coli* (Cirz et al., 2005; Fujikawa et al., 2005). Meanwhile, our wild-type *Mtb* strains were incubated for 40 generations at most in the presence of OFX (see Methods; Table 7.7, see Supplementary Information), making the likelihood of observing OFX-induced OFX-R mutants in our in vitro system extremely low.

Another limitation of our study is that fluctuation analyses only model AMR emergence. Long-term population dynamics also play an important role in AMR evolution (Wiesch et al., 2011; Lindsey et al., 2013; Huseby et al., 2017). For example, population bottleneck events modulate AMR evolution during serial transfer experiments (Comas et al., 2012; Barrick et al., 2013; Vogwill et al., 2016; Huseby et al., 2017), and have also been hypothesized to strongly influence *Mtb* evolution in the clinic (Hershberg et al., 2008). Thus, modeling FQR evolution in *Mtb* in epidemiological settings would benefit from the use of some measure of long-term population dynamics and between-host transmission. Nevertheless, the fitness of AMR mutants is an important factor in determining its evolutionary fate (Wiesch et al., 2011; Angst et al., 2013; Barrick et al., 2013; Lindsey et al., 2013; Hughes et al., 2017; Huseby et al., 2017) and its potential for between-host transmission (Comas et al., 2012; Vos et al., 2013). Considering that the *Mtb* genetic background modulated the fitness effect of FQ-R mutations (Figure 3.5; Table 7.7, see Supplementary Information), the genetic background may modulate how likely FQ-R mutants transmit between patients.

In conclusion, we illustrate how the genetic variation present in natural populations of *Mtb* modulates FQ-R evolution. Considering the nonrandom geographic distribution of different *Mtb* genetic variants (Comas et al., 2010; Gagneux, 2018), our work suggests that there may be regional differences in the rate of FQ-R emergence and FQ-R prevalence when using FQs as a first-line drug. We therefore highlight the importance of standing genetic variation in determining how FQ-R evolves in *Mtb* and, in general, how AMR evolves in pathogens.

4 Simulating the Emergence of Fluoroquinolone-Resistance in *Mycobacterium tuberculosis* using a Mathematical Model

Rhastin A.D. Castro[1,2], Amanda Ross[1,2], Andrej Trauner[1,2], and Sebastien Gagneux[1,2,*]

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland
[2]University of Basel, Basel, Switzerland
[*]Corresponding Author: E-mail: sebastien.gagneux@swisstph.ch.

This Chapter is a working manuscript.

4.1 Abstract

Fluoroquinolones (FQs) are being tested in new experimental treatment regimens against *Mycobacterium tuberculosis* (*Mtb*) infections. In vitro work has shown that the frequency of fluoroquinolone-resistance (FQ-R) emergence in *Mtb* is dependent on bacterial genetics. Further *in vitro* assays suggests that the differences in FQ-R frequencies is not due to *Mtb* strain-dependent mutation rates, but likely due to strain-dependent mutational target sizes for FQ-R and strain-dependent fitness effects of FQ-R mutations. To test the contribution of strain-dependent FQ-R mutational target sizes and FQ-R mutation fitness effects, we adapted the stochastic, time-step model of drug-resistance evolution developed by Ford et al. (2013) to include these parameters. We observed that the use experimentally-measured or inferred strain-dependent values for FQ-R mutational target sizes and FQ-R mutation fitness effects, coupled with the use of the same mutation rate value for every strain, were insufficient to simulate the *Mtb* strain-dependent FQ-R frequencies. Sensitivity analysis showed that while varying the FQ-R mutational target size can modulate FQ-R frequencies *in silico*, varying the mutation rate had the greatest effect. Considering minimal mutation rate differences were observed in vitro, we conclude that other, yet-to-be-identified biological factors that are not included in our model may play a role determining FQ-R frequencies in vitro. Similarly, we show that new models are required to accurately simulate the *Mtb* strain-dependent FQ-R frequencies.

4.2 Introduction

Drug resistance (DR) in pathogens presents both a formidable risk for treatment failure in patients, and an economic burden to patients and health care systems (Laxminarayan et al., 2013; Alvarez-Uria et al., 2016). Multiple socioeconomic and behavioural factors modulate the prevalence of DR (Dalton et al., 2012; Laxminarayan et al., 2013; Eldholm et al., 2016; Alvarez-Uria et al., 2016; Shah et al., 2017). However, the emergence of DR is ultimately an evolutionary process, where an environmental pressure (*i.e.* the antibiotics) selects for a preexisting phenotype in a given population (*i.e.* the DR phenotype) (Luria et al., 1943; Hughes et al., 2017). Studying DR evolution may allow for the development of new strategies that suppress the prevalence of DR in the clinic, potentially saving patient lives and reducing burdens on health care systems.

Previous work suggest that the emergence of DR in pathogens is dependent on multiple biological factors (Wiesch et al., 2011; Hughes et al., 2017). Firstly, DNA mutations are the ultimate source for variation in a population (Luria et al., 1943; Wiesch et al., 2011). Pathogen strains with higher DNA mutation rates have a higher likelihood of acquiring DR mutations, and have also been associated with higher DR prevalence in the clinic (Oliver et al., 2000; Giraud et al., 2001; Gould et al., 2007; Couce et al., 2016). The emergence of DR may also be dependent on the number of potential mutations that can confer DR (hereafter referred to as the mutational target size for DR); pathogens with larger mutational target sizes are hypothesized to associate with higher DR prevalence (Ford et al., 2013; Vog-

will et al., 2014; Hughes et al., 2017). The fitness effects of the DR mutations may also modulate the frequency of DR (Gagneux et al., 2006c; Hughes et al., 2017). Mathematical modeling and molecular epidemiological studies have shown that DR mutations with low or no fitness costs are generally the most prevalent in the clinic (Cohen et al., 2004; Gagneux et al., 2006c; Huseby et al., 2017; Castro et al., 2019). Understanding how these biological factors modulate the rate of DR emergence in pathogens may provide insights into new treatment strategies that could suppress DR emergence and, consequently, restrict DR prevalence.

Pathogen populations comprise genetically distinct strains, and this standing genetic variation has been shown to modulate DR evolution (Gagneux et al., 2006c; Decuypere et al., 2012; Fenner et al., 2012; Apjok et al., 2019; Castro et al., 2019). Of particular note is the impact of pathogen genetics in the evolution of DR in *Mycobacterium tuberculosis* (*Mtb*), the aetiological agent of human tuberculosis (TB). *Mtb* caused the highest number of deaths due to a pathogen in 2018 (WHO, 2019a), and was responsible for nearly 200,000 out of the approximately 700,000 deaths due to a drug-resistant pathogen infection in 2016 (O'Neill, 2016). While the genetic diversity in *Mtb* is relatively small compared to other bacterial pathogens, this limited genetic diversity has substantial consequences on DR evolution (Comas et al., 2010; Gagneux, 2018). For instance, *Mtb* strains belonging to the Lineage 2 (L2) "Beijing" genetic family has repeatedly been associated with multidrug-resistance in the clinic (defined as a patient infected with an *Mtb* strain that is resistant to at least the two most potent first-line drugs, isoniazid and rifampicin) (Borrell et al., 2009; Casali et al., 2014; Merker et al., 2015; Wollenberg et al., 2017). Published molecular epidemiological and in vitro studies suggest that L2 strains may have a "hypermutator" phenotype, which allows for a higher rate of acquiring DR mutations and, consequently, higher DR prevalence (Rad et al., 2003; Ford et al., 2013). However, this hypothesis is disputed, as other published in vitro studies have shown that L2 strains do not acquire DR mutations at higher frequencies (Werngren et al., 2003; Carey et al., 2018). Lastly, bacterial genetics may also affect the frequency of DR emergence in *Mtb* by modulating the mutational target sizes for DR (Gagneux et al., 2006a; Fenner et al., 2012; Ford et al., 2013), and the phenotypic effects of DR mutations (Gagneux et al., 2006c; Zaczek et al., 2009).

In our previous work, we demonstrated that bacterial genetics can modulate the frequency of FQ-R emergence in *Mtb* (Figure 3.1,) (Castro et al., 2019). Specifically, our in vitro results showed minimal mutation rate differences between the different *Mtb* strains tested (Figure 3.2), and suggests that the mutational profiles for FQ-R was the main contributor in leading to the strain-dependent FQ-R frequencies (Figure 3.3). The differences in mutational profiles for FQ-R were driven by the *Mtb* genetic background's effect on both the mutational target size for FQ-R (Figures 3.3 and 3.4), and the fitness effect of FQ-R mutations (Figure 3.5). However, the relative contribution of each parameter in determining the frequency of FQ-R is unclear.

Mathematical models have been used understand or test the relative contributions of biological factors in DR evolution (Spicknall et al., 2013; Blanquart, 2019). For example, theoretical models of DR

evolution have been used to test the role of different drug treatment strategies (Bonhoeffer et al., 1997), the role of competition between the drug-susceptible (DS) and DR populations (Cohen et al., 2004; Hansen et al., 2017; Khan et al., 2018), the role of the immune system (Handel et al., 2009), the role of treatment with sub-inhibitory drug concentrations (Frenoy et al., 2018), and the role of tolerance and persistence phenotypes (Levin-Reisman et al., 2017). For this chapter, we adapted a mathematical model designed by Ford *et al.* (Ford et al., 2013) to test whether the observed *Mtb* strain-dependent FQ-R mutational target sizes and mutation fitness effects would be sufficient in simulating the strain-dependent FQ-R frequencies in vitro (Figure 3.1) (Castro et al., 2019). Further, we tested whether changes to the mutation rate, the mutational target size, or the fitness effects of mutations would lead to the greatest variation in FQ-R frequencies.

4.3 Methods

4.3.1 Model for the Frequency of Fluoroquinolone-resistance in *M. tuberculosis*

We adapted the stochastic, time-step model for the evolution of drug-resistance developed by Ford *et al.* (2013). Our model included the processes of births, deaths, mutation, and the fitness effect from being a mutant, as designed by Ford *et al.* We then extended the model to include multiple types of mutations that are possible for a drug-susceptible cell to mutate into. We used S_{it} to represent the number of drug-susceptible bacteria for genetic background i, where i is one of N0157, N1283, or N0145, and t represents the time-step. Each time-step t was equivalent to 1 hour. The birth rate b represents the rate that new bacteria are added to the population due to cell division per time-step t, while the death rate d represents the rate that bacteria are removed from the population per time-step t due to some form of environmental stress. The parameter m_{ijt} represented the number of bacteria transitioning from the drug-susceptible S_{it} population to the respective drug-resistant R_{ijt} population, where j represents the *gyrA* mutation that the drug-resistant mutant acquired. The *gyrA* mutation j can be one of nine potential FQ-R-conferring *gyrA* mutations observed in the mutational profile assay following a fluctuation analysis of nine *Mtb* strains on 4 μg/ml of the FQ ofloxacin (Figures 3.1 & 3.3; Tables 7.4 & 7.5, see Supplementary Information): G88C, A90V, S91P, D94A, D94G, D94H, D94N, D94Y, or nm (*i.e.* non-*gyrA* mutant, but FQ-R) (Castro et al., 2019). The number of drug-susceptible bacteria at time-step t, S_{it}, is calculated by taking the number of drug-susceptible bacteria at time-step $t-1$, $S_{(it-1)}$, adding new bateria from cell division of drug-susceptible bacteria produced that did not die, $S_{(it-1)} \times (b-d)$, and subtracting the sum of the new *gyrA* mutants, $\sum_{j=1}^{J} m_{ijt}$, so that S_{it} is defined as in Equation (1):

$$S_{it} = S_{(it-1)} + (S_{it-1} \times (b-d)) - \sum_{j=1}^{J} m_{ijt}. \quad (1)$$

The mutation rate μ_i was defined as the probability of a drug-susceptible S_i cell mutating into a drug-resistant R_{ij} cell. As in Ford *et al.* (2013), we set the number of new mutants at time-step t,

m_{ijt}, to follow a Poisson distribution around μ_i multiplied by the drug-susceptible S_i population at time-step $t-1$, so that m_{ijt} is defined as in Equation (2):

$$m_{ijt} \sim \text{Poisson}(\mu_i \times S_{it-1}) \tag{2}$$

The fitness effect of *gyrA* mutation j when present in genetic background i is represented by c_{ij}. A value of 0 for fitness effect c_{ij} meant that the growth rate of the mutant was equivalent to the growth rate of its drug-susceptible ancestor. Fitness effect c_{ij} values above 0 meant that the j mutation was costly, while values below 0 meant that the j mutation conferred a fitness advantage over wild-type.

The number of drug-resistant bacteria with genetic background i and *gyrA* mutation j at time-step t, R_{ijt}, was given by the number of drug-resistant bacteria at time-step $t-1$, R_{ijt-1}, adding new bacteria from cell-division of drug-resistant cells that did not die, $R_{ijt-1}(b(1-c_{ij})-d)$, and adding the new *gyrA* mutants produced at time-step t, m_{ijt}, so that R_{ijt-1} is defined as in Equation (3):

$$R_{ijt} = R_{ijt-1} + (R_{ijt-1}(b(1-c_{ij})-d)) + m_{ijt} \tag{3}$$

Each simulation aimed to simulate the results of an in vitro fluctuation analysis. Thus, every simulation was run 100 times, with each simulation run treated as one parallel culture within an *in silico* fluctuation analysis experiment. At the beginning of every simulation run, the number of drug-susceptible bacteria at time-step 0, S_{i0}, was initially set to a population size of 10. Each simulation ran up to time-step $t = 500$. For every simulation run, we determined the time-step that the total population size (*i.e.* total sum of S_{it} and R_{ijt}) was closest to, but did not exceed 10^9, which we designated as time-step $t = T$. This was an arbitrarily chosen target population size, but a similar value to the bacterial population size that we incubated our in vitro fluctuation analysis cultures to prior to selecting on 4 μg/ml of the FQ ofloxacin (Figure 3.1, see Chapter 3) (Castro et al., 2019). We then calculated the total population size of all drug-resistant mutants at the target population size ($t = T$) per simulation run, as in the Equation (4):

$$\sum_{j=1}^{J} R_{ijt=T} \tag{4}$$

The Ma, Sarkar, Sandri-Maximum Likelihood Estimator method (MSS-MLE) was used to estimate the most likely number of mutations per culture, m, from the distribution of the number of drug-resistant mutants across the 100 simulation runs, r_{dist}, at the time-step $t = T$ (Rosche et al., 2000). The frequency of drug-resistance per cell per strain, F was then calculated by dividing the calculated m values by the total population size (*i.e.* total sum of S_{it} and R_{ijt}) at time-step $t = T$. Significant differences in the frequency of drug-resistance per cell per strain, F, between any two given simulations were defined by non-overlapping 95% confidence intervals (Rosche et al., 2000; Carey et al., 2018). All simulations were performed using the R statistical software (R Core Team, 2018).

4.3.2 Simulation: *In silico* Frequency of Fluoroquinolone-resistance in *M. tuberculosis*

For the first set of simulations, we tested whether we could recapitulate the differential in vitro FQ-R frequencies between the N0157, N1283, and N0145 genetic backgrounds (Figure 3.1, see Chapter 3) using parameter values that were either confirmed or could be inferred from experimental results. These parameter values are listed in Table 4.1. Specifically, we set birth rate b to have a default value of 0.053 for all genetic backgrounds, as this was the average growth rate of the drug-susceptible wild-type strains across the three *Mtb* strains tested in our previous in vitro cell growth assays (N0157, N1283, and N0145), in units of $\log_2(OD_{600})$/hour (Figure 3.5, see Chapter 3; Table 7.7, see Supplementary Information) (Castro et al., 2019). Meanwhile, the death rate d was set to 0 by default, as we were simulating the frequency of drug-resistant mutants in the absence of environmental stress. Considering that we previously observed no hypermutator phenotypes in the N0157, N1283, and N0145 strains (Figure 3.2, see Chapter 3) (Castro et al., 2019), the mutation rate parameter μ_i was set to 4.52×10^{-10} by default; this value was the measured number of mutations per base-pair per generation produced by wild-type *Mtb* DNA polymerase DnaE1 in a mutation accumulation assay performed by Rock *et al.* (2015). In the absence of stress, chromosomal mutational events generally occur during DNA replication events and, consequently, should only happen with every cell division event rather than at every hour. Thus, mutation rate μ_i was corrected for by dividing it by the generation time of drug-susceptible S_i. Considering that the birth rate b was equivalent to the $\log_2(OD_{600})$/hour of drug-susceptible S_i, the generation time of drug-susceptible S_i was equivalent to $1/b$.

Table 4.1: Parameter values used for simulating the frequency of fluoroquinolone-resistance in the N0157, N1283, and N0145 strains

Parameter	N0157 simulation	N1283 simulation	N0145 simulation
b	0.053	0.053	0.053
d	0	0	0
μ	4.52×10^{-10}	4.52×10^{-10}	4.52×10^{-10}
c_{G88C}	0.156	0.311	1
c_{A90V}	-0.156	0.115	1
c_{S91P}	1	0.311	1
c_{D94A}	1	1	1
c_{D94G}	0.0222	-0.0164	0.0566
c_{D94H}	0.156	0.295	1
c_{D94N}	0.0667	0.148	0.0377
c_{D94Y}	0.156	0.148	1
c_{nm}	0.156	1	1

The fitness effect c_{ij} was mutation- and genetic background-specific, and set based on measured values in our previous in vitro cell growth assays (Figure 3.5, see Chapter 3; Table 7.7, see Supplementary Material) (Castro et al., 2019). In cases where no cell growth assay measurements were present, we inferred the cost of the mutation based on its relative frequency in the mutational profile assay follow-

ing a fluctuation analysis of nine *Mtb* strains on 4 μg/ml of ofloxacin (Figures 3.1 & 3.3, see Chapter 3; Tables 7.4 & 7.5, see Supplementary Information) (Castro et al., 2019). For example, while no growth rate measurements were available for the GyrA D94H mutation when present in the N0157 genetic background, it was found in the same frequency as the G88C mutation in the mutational profile assay; therefore, the c_{ij} parameter value for GyrA D94H was set to be equivalent to the GyrA G88C mutation when present in the N0157. Another example is GyrA S91P mutation, which was not present in the mutational profile assay for N0157; thus its fitness effect c_{ij} value was set to 1.00 (fully lethal) for the N0157 simulations.

4.3.3 Simulation: Sensitivity Analysis

We performed a sensitivity analysis of our model to test whether the mutation rate, the mutational target size for FQ-R, or the fitness effects of FQ-R mutations would have the greatest impact on the *in silico* frequency of FQ-R. To do so, we ran different sets of simulations where we let one parameter's value vary while keeping the other two parameters' values fixed. We ran an initial simulation (designated as simulation SA.1) to compare the other test simulations against. The SA.1 simulation used a default set of parameter values, with a mutation rate μ_i value of 4.52×10^{-10}. We set the default mutational target size to equal 8 mutations in SA.1 by setting the "nm" mutation to be lethal ($c_{\mathrm{nm}} = 1.00$). This was done so that the values set for each parameter can be made different from their SA.1 value by a multiplicative of 2 in subsequent simulations. Each of the 8 FQ-R mutations that are possible in SA.1 had their fitness cost set at 5% (c_{ij} = 0.05). To test the impact of mutation rates on the frequency of FQ-R *in silico*, we ran two simulations where the mutation rate value was equal to either one-half (designated as SA.2) or one-quarter (SA.3) of the mutation rate value in SA.1. We next tested the impact of fitness costs in a similar manner by running two simulations that had double (SA.4) or quadruple (SA.5) the fitness cost values for each of 8 potential FQ-R mutations compared to their SA.1 values. Finally, we tested the impact of mutational target sizes using two simulations where the mutational target size was reduced to 4 (SA.6) or to 2 (SA.7). The parameter values for all seven Sensitivity Analysis simulations are presented in Table 4.2.

Table 4.2: Parameter values used for the sensitivity analysis of the model for the emergence of fluoroquinolone-resistance

Param.	SA.1	SA.2	SA.3	SA.4	SA.5	SA.6	SA.7
b	0.053	0.053	0.053	0.053	0.053	0.053	0.053
d	0	0	0	0	0	0	0
μ	4.52×10^{-10}	2.26×10^{-10}	1.13×10^{-10}	4.52×10^{-10}	4.52×10^{-10}	4.52×10^{-10}	4.52×10^{-10}
c_{G88C}	0.05	0.05	0.05	0.10	0.20	1	1
c_{A90V}	0.05	0.05	0.05	0.10	0.20	0.05	0.05
c_{S91P}	0.05	0.05	0.05	0.10	0.20	1	1
c_{D94A}	0.05	0.05	0.05	0.10	0.20	1	1
c_{D94G}	0.05	0.05	0.05	0.10	0.20	0.05	0.05
c_{D94H}	0.05	0.05	0.05	0.10	0.20	1	1
c_{D94N}	0.05	0.05	0.05	0.10	0.20	0.05	1
c_{D94Y}	0.05	0.05	0.05	0.10	0.20	0.05	1
c_{nm}	1	1	1	1	1	1	1

4.3.4 Simulation: Simulating the *M. tuberculosis* strain-specific Differences in the Frequency of Fluoroquinolone-resistance In Vitro

To simulate the *Mtb* strain-specific differences in the frequency of FQ-R in vitro, we again used the fitness costs and mutational target size values for FQ-R that were either confirmed or could be inferred from experimental results (Figures 3.3 & 3.5, see Chapter 3; Tables 7.4, 7.5, & 7.7, see Supplementary Information) (Castro et al., 2019). Thus, the fitness effects of FQ-R mutations c_{ij} were identical to those used in the set of simulations simulating the frequency of FQ-resistance when using the same mutation rate μ_i value of 4.52×10^{-10} for the N0157, N1283, and N0145 strains (Table 4.1). However, for these sets of simulations, the mutation rate μ_i was made variable for each simulation and for every strain. All parameters used are listed in Table 4.3. Specifically, for the N0157 simulations, 100 different mutation rate μ_i values were tested. These ranged from a minimum value of 3.55×10^{-10} to a maximum value of 3.55×10^{-8}, and all values were equidistant from each other in magnitude. For the N1283 simulations, the mutation rate μ_i values tested ranged from 1.98×10^{-11} to 1.98×10^{-9}, while for the N0145 simulations, the mutation rate μ_i values ranged from 1.53×10^{-12} to 1.53×10^{-10}. These mutation rate μ_i value ranges were chosen as they produced a range of values for the frequency of FQ-R mutations per cell (as calculated by MSS-MLE) that were similar to what was observed in the in vitro fluctuation analysis on 4 μg/ml ofloxacin (Figure 3.1, see Chapter 3) (Castro et al., 2019). For each *Mtb* strain, we defined the model's estimate of the strain's mutation rate as the initial mutation rate μ_i value set in a simulation that resulted in a frequency of FQ-R mutations per cell, F, that was closest to the in vitro fluctuation analysis on 4 μg/ml ofloxacin (Figure 3.1, see Chapter 3) (Castro et al., 2019).

Table 4.3: Parameter values used to simulate the strain-specific differences in the frequency of fluoroquinolone-resistance in vitro

Parameter	N0157 simulation	N1283 simulation	N0145 simulation
b	0.053	0.053	0.053
d	0	0	0
μ_{min}	3.55×10^{-10}	1.98×10^{-11}	1.53×10^{-12}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
μ_{max}	3.55×10^{-8}	1.98×10^{-9}	1.53×10^{-10}
c_{G88C}	0.156	0.311	1
c_{A90V}	-0.156	0.115	1
c_{S91P}	1	0.311	1
c_{D94A}	1	1	1
c_{D94G}	0.0222	-0.0164	0.0566
c_{D94H}	0.156	0.295	1
c_{D94N}	0.0667	0.148	0.0377
c_{D94Y}	0.156	0.148	1
c_{nm}	0.156	1	1

There are 98 μ_i values between μ_{min} and μ_{max}, with each value equidistant in magnitude from each other, from μ_{min}, and from μ_{max}.

4.4 Results

4.4.1 *In silico* Frequency of Fluoroquinolone-resistance in *M. tuberculosis*

Based on our previous experimental data (Castro et al., 2019), the mutational target size for FQ-R and the fitness effects of FQ-R mutations may play a role in determining FQ-R frequencies in *Mtb*. To explore their relative contributions, we adapted the stochastic time-step model on the evolution of DR developed by Ford *et al.* (2013) to simulate the frequency of FQ-R in *Mtb*. As in Ford *et al.*, our model included the processes of births, deaths, mutation, and the fitness effect from being a drug-resistant mutant (Figure 4.1). However, we extended the model so that a drug-susceptible cell can mutate and become an FQ-R mutant with one of the nine possible mutations observed in the mutational profile assay following a fluctuation analysis on 4 μg/ml of OFX (Figures 3.1 & 3.3, see Chapter 3; Tables 7.4 & 7.5, see Supplementary Information) (Castro et al., 2019). We used this model to simulate the population dynamics of an in vitro fluctuation analysis for FQ-R for three strains with differential FQ-R frequencies in vitro: N0157 (high FQ-R frequency), N1283 (mid-FQ-R frequency), and N0145 (low FQ-R frequency). We used experimentally-determined values for the FQ-R mutational target sizes and fitness costs for each *Mtb* strain when we had the relevant measurement, and inferred parameter values if they were missing (see Methods). These simulation values are presented in Table 4.1. We expected that the strain-dependent variation in both the mutational target sizes and relative fitness costs for FQ-R would be sufficient to simulate the strain-dependent in vitro FQ-R frequencies variation (Figure 3.1, see Chapter 3) (Castro et al., 2019). However, this was not the case (Figure 4.2;

Table 4.4). We observed that the frequency of FQ-R acquired per cell (F) for the N0157 simulation (designated as N0157_S) was one order of magnitude lower than what was observed following the in vitro fluctuation analysis (designated as N0157_F) ($F_{\text{N0157\_S}} = 3.26 \times 10^{-9}$, 95% CIs $[2.74 \times 10^{-9}$, $3.81 \times 10^{-9}]$; $F_{\text{N0157\_F}} = 3.55 \times 10^{-8}$, 95% CIs $[2.75 \times 10^{-8}, 4.42 \times 10^{-8}]$). Conversely, the simulation for N0145 (N0145_S) produced an F value that was one order of magnitude higher than what was observed in the in vitro fluctuation analysis (N0145_F) ($F_{\text{N0145\_S}} = 2.41 \times 10^{-9}$, 95% CIs $[1.99 \times 10^{-9}, 2.86 \times 10^{-9}]$; $F_{\text{N0145\_F}} = 1.53 \times 10^{-10}$, 95% CIs $[5.44 \times 10^{-11}, 2.85 \times 10^{-10}]$). The N1283 simulation (N1283_S) was the only simulation that provided an F value that did not differ significantly from the observed in vitro value (N1283_F) based on their overlapping 95% confidence intervals ($F_{\text{N1283\_S}} = 3.02 \times 10^{-9}$, 95% CIs $[2.53 \times 10^{-9}, 3.55 \times 10^{-9}]$; $F_{\text{N1283\_F}} = 1.98 \times 10^{-9}$, 95% CIs $[1.35 \times 10^{-9}, 2.70 \times 10^{-9}]$). Thus, our results suggests that the mutational target size and the fitness costs for FQ-R do not have large contributions in determining the frequency of FQ-R in *Mtb* in our *in silico* model. We tested this hypothesis in the next section.

Table 4.4: Frequency of fluoroquinolone-resistance per cell following an in vitro fluctuation analysis at 4 μg/ml of the FQ ofloxacin and an *in silico* simulation

Strain	Experiment	F	F: Lower 95% CI	F: Upper 95% CI	Mean Population Size
N0157	N0157_F	3.55×10^{-8}	2.75×10^{-8}	4.42×10^{-8}	2.00×10^{8}
N0157	N0157_S	3.26×10^{-9}	2.74×10^{-9}	3.81×10^{-9}	1.05×10^{9}
N1283	N1283_F	1.98×10^{-9}	1.35×10^{-9}	2.70×10^{-9}	1.10×10^{9}
N1283	N1283_S	3.02×10^{-9}	2.53×10^{-9}	3.55×10^{-9}	1.02×10^{9}
N0145	N0145_F	1.53×10^{-10}	5.44×10^{-11}	2.85×10^{-10}	9.57×10^{8}
N0145	N0145_S	2.41×10^{-9}	1.99×10^{-9}	2.86×10^{-9}	1.02×10^{9}

Frequency of fluoroquinolone-resistance (F) per cell was reported for three *M. tuberculosis* strains (N0157, N1283, and N0145) following in vitro fluctuation analysis at 4 μg/ml of ofloxacin (denoted by the suffix "_F" in the Experiment column) or following the *in silico* simulation (denoted by the suffix "_S").

4.4.2 Sensitivity Analysis

Using a sensitivity analysis experiment, we tested whether the bacterial mutation rate, the mutational target size for FQ-R, or the fitness effects of FQ-R mutations had the greatest effect on the frequency of FQ-R *in silico*. An initial simulation was run (designated as simulation SA.1), where a mutation rate μ_i value of 4.52×10^{-10} was used (as measured by Rock et al., 2015) with a mutational target size of 8 mutations. All 8 potential mutations had fitness costs of 5% (*i.e.* $c_{ij} = 0.05$). Using these parameters, we observed that the frequency of FQ-R acquired per cell for SA.1, $F_{\text{SA.1}}$, was equal to 3.68×10^{-9} (95% CIs $[3.12 \times 10^{-9}, 4.28 \times 10^{-9}]$) (Figure 4.3; Table 4.5). When we simulated different input mutation rate μ_i values, we observed that dividing the mutation rate μ_i value from SA.1 by 2 (SA.2) or by 4 (SA.3) both led to a significant decreases in the observed frequency of FQ-R

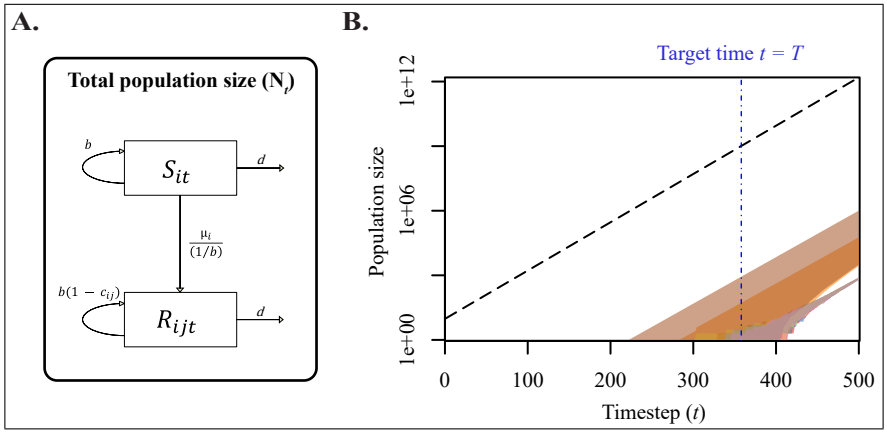


Figure 4.1: Visualization of the stochastic model for the frequency of fluoroquinolone-resistance. A. The structure of the model. The population size of drug-susceptible and drug-resistant bacteria at time-step t are represented by S_{it} and R_{ijt}, respectively. Total population size N_t is defined by taking the sum of S_{it} and R_{ijt}. The bacterial birth and death rates per time-step t are represented by b and d, respectively. The fitness effect of *gyrA* mutant j is represented by c_{ij}, with positive values conferring a fitness cost and negative values conferring a fitness gain. The mutation rate per time-step t is represented by μ_i. So that mutations occur per cell division and not per time-step t, the mutation rate μ_i is corrected for dividing it by the generation time of drug-susceptible S_i, which is equivalent to $1/b$. B. Population size for sensitive and resistant bacteria across 100 simulation runs testing the frequency of fluoroquinolone-resistance. The results for N0145 are used as an example. Dashed black line represents population size of drug-susceptible S_{N0145} over time. The variability in the S_{N0145} across 100 simulation runs was negligible, as the population lost due to m_{ijt} in Equation (2) was extremely small compared to the total S_{N0145}. The variability in the population sizes of each *gyrA* mutant population ($R_{N0145,j}$) across all simulation runs are visualized as coloured polygons, with each colour denoting a different *gyrA* mutant j. The upper-bound of the coloured polygons show the highest population size, while the lower-bound shows the lowest population size, for that given mutant across all simulation runs. The time-point t that the total population size was closest to but did not exceed 1×10^9 was identified (Target $t = T$, in blue dash-dotted line). The total drug-resistant population ($R_{N0145,j}$) at target time-point $t = T$ was then determined for each simulation to create r_{dist}. The frequency of drug-resistance per cell was then calculated from r_{dist} using the MSS-MLE method (Rosche et al., 2000).

compared to SA.1 ($F_{\text{SA.2}} = 2.03 \times 10^{-9}$, 95% CIs $[1.66 \times 10^{-9}, 2.43 \times 10^{-9}]$; $F_{\text{SA.3}} = 1.01 \times 10^{-8}$, 95% CIs $[7.83 \times 10^{-10}, 1.26 \times 10^{-9}]$). Conversely, increasing the fitness costs to double (SA.4) or quadruple (SA.5) the values from SA.1 for every FQ-R mutation had no significant effects on the frequency of FQ-R compared to SA.1 ($F_{\text{SA.4}} = 3.60 \times 10^{-9}$, 95% CIs $[3.05 \times 10^{-9}, 4.19 \times 10^{-9}]$; $F_{\text{SA.5}} = 3.74 \times 10^{-9}$, 95% CIs $[3.17 \times 10^{-9}, 4.35 \times 10^{-9}]$). Reducing the mutational target size from 8 to 4 (SA.6) had no significant effect on FQ-R frequency FQ-R ($F_{\text{SA.6}} = 2.88 \times 10^{-9}$, 95% CIs $[2.40 \times 10^{-9}, 3.38 \times 10^{-9}]$), but further reducing the mutational target size down to 2 (SA.7) did lead to a significant decrease compared to SA.1 ($F_{\text{SA.7}} = 2.50 \times 10^{-9}$, 95% CIs $[2.07 \times 10^{-9}$, $2.96 \times 10^{-9}]$). However, the four-fold reduction in the mutational target size (SA.7) led to a smaller decrease in the frequency of FQ-R than what was observed following a four-fold reduction in the mu-

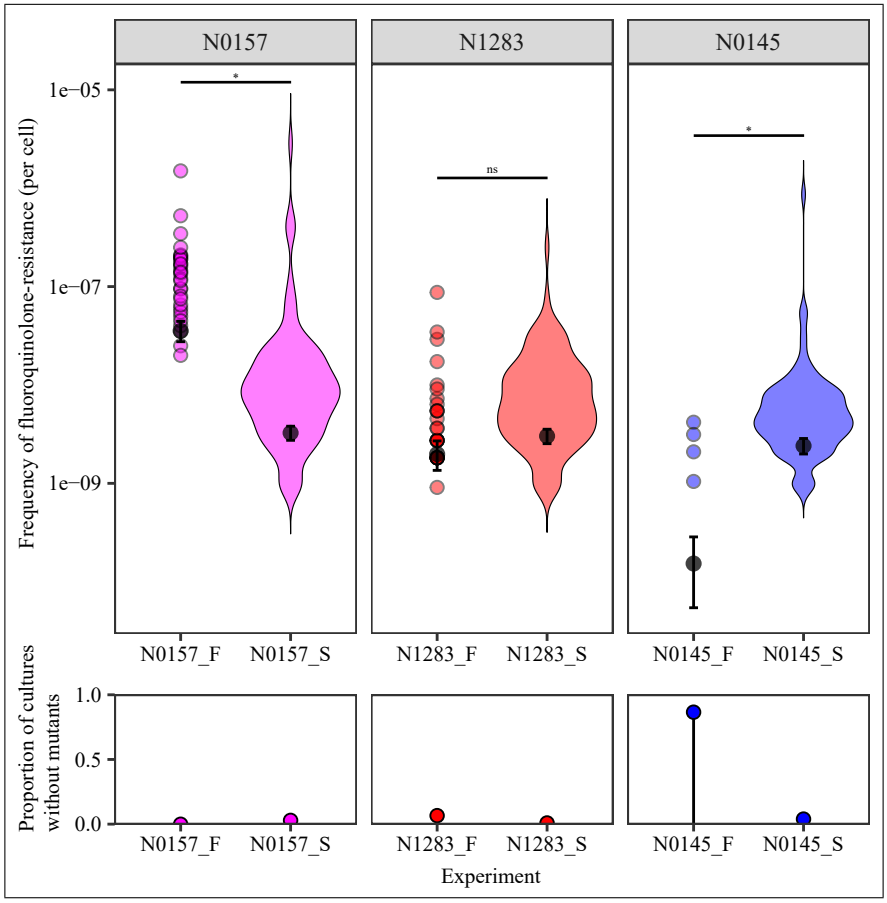


Figure 4.2: Frequency of fluoroquinolone-resistance in vitro versus *in silico* for three *M. tuberculosis* strains.
Top panels: Coloured points represent frequency of fluoroquinolone-resistance per cell per parallel culture as measured by fluctuation analysis at 4 μg of ofloxacin (suffixed with "_F" after *M. tuberculosis* strain names N0157, N1283, and N0145), with darker points representing multiple cultures with the same frequency. Fluctuation analysis data presented here is identical as in Figure 3.1 (see Chapter 3) (Castro et al., 2019). Colours denote the lineage that the *M. tuberculosis* strain belongs to (L1 = pink; L2 = blue; L4 = red). Violin plots show the distribution of the frequency of resistance per cell across 100 simulated cultures (suffixed with "_S"), per genetic background. Grey points represent the estimated number of mutations per cell across all parallel cultures as calculated by MSS-MLE (Rosche et al., 2000), while black bars denote the respective 95% confidence intervals. Significant differences in the frequency of fluoroquinolone-resistance per cell between the in vitro fluctuation analysis data and their respective simulation counterparts were defined by non-overlapping 95% confidence intervals (Rosche et al., 2000; Carey et al., 2018). Bottom panels: the percentage of parallel cultures lacking drug-resistant mutants.

tation rate (SA.3) when compared to SA.1. These results suggest that the mutation rate had a greater impact on the *in silico* frequency of FQ-R compared to either the FQ-R mutational target size or fitness effects of FQ-R mutations in our model.

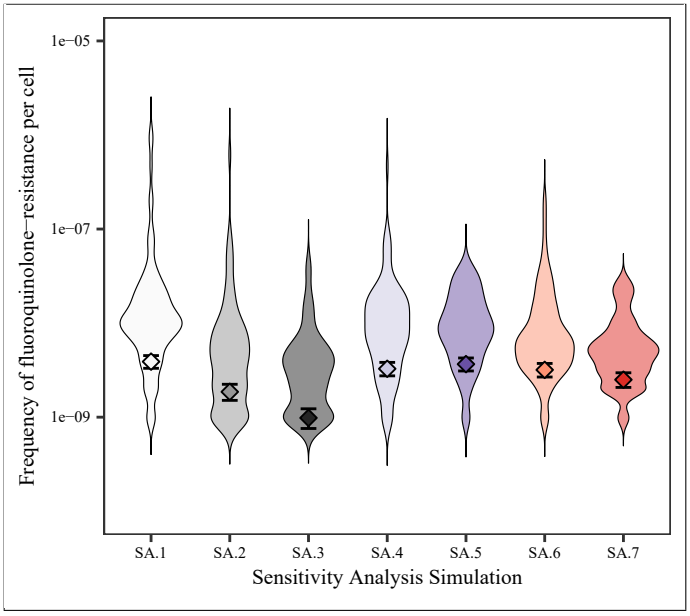


Figure 4.3: Sensitivity analysis of the model for the frequency of fluoroquinolone-resistance, using simulation parameters listed in Table 4.5.
Violin plot in white (left-most plot; SA.1) shows the r_{dist} for the simulation using default parameters, with the diamond representing the frequency of resistance per cell as calculated by MSS-MLE (Rosche et al., 2000). Violin plots and diamonds coloured in different shades of grey show the r_{dist} for simulations using variable mutation rates μ_i (SA.2 & SA.3). Violin plots and diamonds in different shades of purple show simulation results using variable fitness effect parameters c_{ij} (SA.4 & SA.5), while plots in different shades of red show simulations with variable mutational target sizes (SA.6 & SA.7).

Table 4.5: Frequency of fluoroquinolone-resistance per cell per Sensitivity Analysis (SA) simulation.

Simulation	F	F: Lower 95% CI	F: Upper 95% CI
SA.1	3.68×10^{-9}	3.12×10^{-9}	4.28×10^{-9}
SA.2	2.03×10^{-9}	1.66×10^{-9}	2.43×10^{-9}
SA.3	1.01×10^{-9}	7.83×10^{-10}	1.26×10^{-9}
SA.4	3.60×10^{-9}	3.05×10^{-9}	4.19×10^{-9}
SA.5	3.74×10^{-9}	3.17×10^{-9}	4.35×10^{-9}
SA.6	2.88×10^{-9}	2.40×10^{-9}	3.38×10^{-9}
SA.7	2.50×10^{-9}	2.07×10^{-9}	2.96×10^{-9}

Parameters for each SA simulation are presented in Table 4.2.

4.4.3 Simulating the *M. tuberculosis* Strain-specific Differences in the Frequency of Ofloxacin-resistance In Vitro

We previously showed that we cannot simulate the *Mtb* strain-specific frequency of FQ-R in vitro while using the same mutation rate μ_i value per strain (Figure 4.2; Table 4.4). Indeed, our sensitivity analysis showed that the bacterial mutation rate had the greatest impact on the *in silico* frequency of

FQ-R in our model (Figure 4.3; Table 4.5). Therefore, we next aimed to explicitly simulate the strain-specific frequency of FQ-R in vitro by using different mutation rate μ_i values per strain. We again used the strain-specific fitness costs and mutational target size values for FQ-R that were either confirmed or could be inferred from experimental results; however, for these set of simulations, the mutation rate μ_i values were also made variable per simulation per strain. All parameter values are listed in Table 4.3.

We found that in order to simulate the observed differences in the in vitro frequency of FQ-R between N0157, N1283, and N0145, the input mutation rate μ_i values per strain had to be different (Table 4.6). Specifically, the model broadly estimated that the N0157 strain would have to have had a mutation rate of 4.62×10^{-9}, the N1283 strain mutation rate would be 2.78×10^{-10}, while the N0145 strain mutation rate would be 1.68×10^{-11} to simulate their differences in FQ-R frequency in vitro. The relative difference in the estimated mutation rates between the different *Mtb* strains (one order of magnitude difference between N0157 and N1283, one order of magnitude difference between N1283 and N0145, and two orders of magnitude difference between N0157 and N0145) mirrored the relative differences in the frequency of FQ-R mutations per cell in vitro (Figure 3.1, see Chapter 3; Table 4.4) (Castro et al., 2019).

Therefore, our model suggested that there may be large differences in mutation rates between our *Mtb* strains. Alternatively, a yet-to-be identified biological factor or factors may also be contributing significantly to the *Mtb* strain-dependent variation in FQ-R frequency in vitro.

Table 4.6: Mutation rate estimates for the N0157, N1283, and N0145 *Mtb* strains based on the comparison between the simulation for fluoroquinolone-resistance emergence and the in vitro fluctuation analyses on 4 μg/mL of ofloxacin.

	N0157 Simulation	N1283 Simulation	N0145 Simulation
Model estimate of mutation rate*	4.62×10^{-9}	2.78×10^{-10}	1.68×10^{-11}

*in number of DNA mutations per cell per generation.

4.5 Discussion

Mathematical models provide a useful method to understand the relative contributions of different biological factors in DR evolution (Spicknall et al., 2013; Blanquart, 2019). In this Chapter, we used the stochastic time-step model developed by Ford *et al.* (2013) to test the roles of the bacterial mutation rate, the mutational target size for FQ-R, and the fitness effects of FQ-R mutations in determining the frequency of FQ-R from an initially drug-susceptible population of *Mtb*. While our previous in vitro work suggests that the mutational target size for FQ-R and the fitness effects of FQ-R mutations are the greatest contributor in determining the strain-dependent frequency of FQ-R in *Mtb* (Castro et al., 2019), our *in silico* results suggest that mutation rates play the greatest role. However, our previous

in vitro results showed little to no strain-dependent variation in mutation rates (Castro et al., 2019). While we included into our model all biological parameters that are hypothesized to be relevant in the emergence of drug-resistance mutations in the absence of external stress or competition (Ford et al., 2013; Hughes et al., 2017), the mismatch between our *in silico* and in vitro results implies that important parameters were missing. Three potential parameters are discussed below.

First, the rate of cell death may play a role in the frequency of FQ-R in *Mtb*. An appreciable rate of cell death has been shown to lead to an overabundance of mutations under stressful conditions and, consequently, to an overestimation of mutation rates (Frenoy et al., 2018). If cell death rates did play a role in our study, then *Mtb* strains with higher rates of cell death during our fluctuation analyses would have higher frequencies of FQ-R. While we did not explicitly measure death rates in our fluctuation analyses, potential differences in death rates are likely negligible in our in vitro system as we incubated our cultures under stress-free and nutrient-rich conditions.

Secondly, differences in the rate of acquiring phenotypic mutations may influence the frequency of FQ-R, and is also not considered in our model. Phenotypic mutations occur during transcription, translation, epigenetic modifications, or due to protein promiscuity, ultimately leading to non-heritable errors in protein structure and function (Loftfield et al., 1972; Ozbudak et al., 2002; Payne et al., 2019b). Phenotypic mutations can lead to differences in antimicrobial tolerance and persistence, which may potentiate the acquisition of heritable DNA mutations for DR (Javid et al., 2014; Levin-Reisman et al., 2017; Hicks et al., 2018). Our panel of nine *Mtb* strains may have differences in their rate of phenotypic mutations acquisition, and *Mtb* strains with higher phenotypic mutation rates may have higher FQ-R frequencies.

Lastly, our model did not include efflux pump expression as a parameter. While the clinical relevance of efflux pump overexpression for FQ-R in *Mtb* has not been established, it has been observed in vitro (Gygli et al., 2017). Our previous genomic analysis of the nine *Mtb* strains used in the fluctuation analysis at 4 μg/ml ofloxacin showed no mutations in genes that putatively control efflux pump expression (Castro et al., 2019). However, there may be strain-dependent differences in efflux pump expression when exposed to FQ, and we did not test for this possibility. As with phenotypic mutations, up-regulation of efflux pumps expression may potentiate the acquisition of heritable DNA mutations for drug-resistance.

Although our model failed to capture the in vitro frequency of FQ-R in the N0157 and N0145 strains when using the relevant FQ-R parameter values from our previous in vitro work (Castro et al., 2019), the model accurately simulated the frequency of FQ-R for the N1283 strain. Considering that approximately half of the *Mtb* strains in our panel had a similar FQ-R frequency as N1283 (Figure 3.1, See Chapter 3) (Castro et al., 2019), we conclude that our model can simulate the in vitro evolution of FQ-R for many strains of *Mtb*. Our results also suggest that more work is required to elucidate all relevant biological mechanisms that leads to the *Mtb* strain-dependent differences in FQ-R frequencies. Similarly, the design of new DR evolution models for *Mtb* are required.

5 Testing the Impact of Fluoroquinolone-resistance on the Genetic Diversity of *Mycobacterium tuberculosis*

Rhastin A.D. Castro[1,2], Chloé Loiseau[1,2], Lujeko Kamwela[1,2], Miriam Reinhard[1,2], Julia Feldmann[1,2], Sonia Borrell[1,2], Daniela Brites[1,2], Andrej Trauner[1,2], and Sebastien Gagneux[1,2,*]

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland

[2]University of Basel, Basel, Switzerland

[*]Corresponding Author: E-mail: sebastien.gagneux@swisstph.ch.

This Chapter is a working manuscript.

5.1 Abstract

Fluoroquinolones (FQs) are a critical component in experimental treatment regimens against drug-susceptible tuberculosis (TB), and in current treatment regimens against multidrug-resistant TB (MDR-TB). Investigating how populations of *Mycobacterium tuberculosis* (*Mtb*; the etiological agent of TB) evolve under FQ pressure may provide insights into maintaining the current potency and future clinical use of FQs. Clinically-relevant FQ-resistance (FQ-R) mutations are restricted to chromosomal mutations in genes encoding DNA gyrase: *gyrA* and *gyrB*. FQ-R mutations may disrupt the DNA gyrase function of regulating DNA topology, leading to increased DNA damage; this may induce hypermutator phenotypes, and increase the adaptive potential of *Mtb* populations through increased genetic diversity. Therefore, we tested whether FQ-R mutations can increase frequencies of acquiring further drug-resistance mutations. We observed that FQ-R can associate with increased frequencies of streptomycin-resistance (STR-R) in vitro; this associated increase in STR-R was dependent on both the FQ-R *gyrA* mutation present, as well as in which *Mtb* strain the *gyrA* mutation was present. We then tested whether FQ-R mutations associated with increased genetic diversity in natural populations of *Mtb*. Analysis of 2,502 publicly available *Mtb* genomes isolated from a global sample of MDR-TB patients, as well as *Mtb* genomes isolated from 118 serially sampled MDR-TB patients from Georgia, showed no evidence for an association between FQ-R mutations and increased genetic diversity. We therefore conclude that FQ-R mutations do not induce stable nor transient hypermutator phenotypes in *Mtb*. Furthermore, we demonstrate the usefulness of evolutionary studies in predicting how bacterial populations respond to new treatment regimens.

5.2 Introduction

Tuberculosis (TB) caused the most deaths in humans due to a single infectious agent in 2018, with an estimated 1.2 million deaths due to TB alone and an additional 0.3 million deaths due to TB-HIV co-infections (WHO, 2019a). This is despite the availability of antimicrobials that can be used to treat infections caused by *Mycobacterium tuberculosis* (*Mtb*), the etiological agent of TB. Two of the most potent antimicrobials against *Mtb*, isoniazid (INH) and rifampicin (RIF), are an integral part of first-line treatment regimen against drug-susceptible strains. However, this first-line treatment regimen requires taking daily doses for 6 to 9 months. This long-term and substantial treatment regimen can be a burden to healthcare systems and can have low patient adherence (Munro et al., 2007; Barter et al., 2012; Alipanah et al., 2018; Ruru et al., 2018). Further compounding the difficulty in TB treatment is the rise of multidrug-resistant TB (MDR-TB), defined as an infection with an *Mtb* strain that is resistant to at least RIF and INH. Treatment for MDR-TB can be 9 to 24 months in duration, and generally have lower treatment success rates than first-line treatment regimen (Kibret et al., 2017; Parmar et al., 2018; Zhang et al., 2018; WHO, 2019a). Thus, there is a vital need to design and produce shorter and more potent treatment regimens that treat both fully drug-susceptible and drug-resistant

strains of *Mtb*.

Fluoroquinolones (FQs) are currently being tested to be part of new treatment regimens against drug-susceptible and drug-resistant strains of *Mtb* (Gillespie et al., 2014; Jindani et al., 2014; Merle et al., 2014; Imperial et al., 2018; Vjecha et al., 2018). FQs have been and are currently used to treat MDR-TB patients (Takiff et al., 2011; Pranger et al., 2019). The previous use of FQs has led to the development of FQ-R in *Mtb* populations, with FQ-R being one of the defining properties of extensively drug-resistant TB (XDR-TB). However, newer generations FQs are more potent against *Mtb* compared to previous generations (Ji et al., 1995; Ji et al., 1998; Yu et al., 2016). These new generation of FQs are associated with greater treatment success rates in *Mtb*-infected mice and humans (Nuermberger et al., 2004; Ahmad et al., 2018), and may even be used against some *Mtb* strains that were resistant to the older generations of FQs (Chien et al., 2016; Maitre et al., 2017; Pranger et al., 2019). Consequently, one of the aims for the experimental first-line regimens containing FQs is to decrease treatment duration down to 4 months for drug-susceptible *Mtb*, and down to 6 months for MDR-TB. FQs are therefore a critical component in the current and potential strategies to treat TB. Investigations into how *Mtb* populations evolve to FQs pressure may aid in maintaining the potency and potential use of FQs.

The sole target of FQs in *Mtb* is DNA gyrase (Takiff et al., 1994; Piton et al., 2010; Maruri et al., 2012). DNA gyrase is the only type II topoisomerase in *Mtb* (Cole et al., 1998) and is responsible for introducing negative supercoils in the bacterial chromosome (Gellert et al., 1976; Levine et al., 1998). DNA gyrase achieves this through an ATP-dependent process that involves creating a double-stranded DNA break (DSB), looping a segment of the chromosome through the DSB, and then re-ligating the DSB (Gellert et al., 1976; Levine et al., 1998). Therefore, DNA gyrase is essential for regulating DNA topology, transcription, and DNA replication (Levine et al., 1998). DNA gyrase is composed of two subunits of GyrA and two subunits of GyrB, which are encoded by the *gyrA* and *gyrB*, respectively (Cole et al., 1998; Levine et al., 1998). The region where FQs binds to in DNA gyrase, known as the "Quinolone-binding Pocket" (QBP), is located within the active site of the enzyme (Piton et al., 2010; Aldred et al., 2016; Blower et al., 2016); consequently, the presence of a FQ molecule within the QBP inhibits the re-ligation activity of DNA gyrase, leading to the toxic effect of FQs (Piton et al., 2010; Aldred et al., 2016; Blower et al., 2016). Clinically-relevant FQ-R mutations are restricted to a limited set of mutations in the "quinolone-resistance-determining region" (QRDR) of *gyrA* and *gyrB* (Takiff et al., 1994; Maruri et al., 2012; Castro et al., 2019). The QRDR is also located in the same region as the QBP, showing a direct relationship between the FQ-DNA gyrase binding site and FQ-R mutations (Piton et al., 2010; Aldred et al., 2016; Blower et al., 2016). Mutations in the QRDR can affect DNA gyrase structure. Structural characterization and molecular dynamics-based simulations suggest that the two most prevalent FQ-R mutations, *gyrA* A90V and D94G (Maruri et al., 2012; Castro et al., 2019), reduce FQ-binding either by disrupting the biochemical interaction between FQs and DNA gyrase (Aldred et al., 2016; Blower et al., 2016), or by increasing the surface area and volume of the QBP and leading to a decreased stability of FQs within the QBP (Pandey et al., 2018).

Structural changes caused by QRDR mutations may affect not only FQ-R binding, but the efficiency of DNA gyrase supercoiling functions as well. For example, we hypothesize that structural changes in the QBP may affect the precision of the DNA gyrase's re-ligation of DSB. This would consequently result in DNA damage, and the presence of either single-stranded DNA breaks or DSBs would lead to the activation of DNA repair mechanisms (Ayora et al., 2011; Singh, 2017). These DNA repair mechanisms may be error-prone, leading to changes in the DNA sequence and higher mutation rates observed (Gong et al., 2005; Aniukwu et al., 2008; Gupta et al., 2011). The use of FQs against *Mtb* infections may therefore lead to FQ-R mutant strains of *Mtb* that also have a hypermutator phenotype.

Hypermutator phenotypes may increase the adaptive capabilities of bacterial populations (Taddei et al., 1997; Hughes et al., 2017). Namely, due to the greater mutation supply generated by hypermutator phenotypes, rare beneficial mutations have a higher likelihood of emerging in bacterial populations with hypermutators (Giraud et al., 2001; Wielgoss et al., 2013; Lynch et al., 2016; Hughes et al., 2017; Raynes et al., 2018). Indeed, hypermutator phenotypes have been associated with increased drug-resistance (DR) mutations in clinical populations of *Pseudomonas aeruginosa* (Oliver et al., 2000; Maciá et al., 2005), *Escherichia coli* (Örlén et al., 2006), and *Staphylococcus aureus* (Prunier et al., 2003; Wang et al., 2013). However, hypermutator phenotypes also increase the accumulation of common deleterious mutations, a phenomenon known as genetic load (Haldane, 1937; Muller, 1950; Zhang et al., 2019). Thus, hypermutator phenotypes may be transient in a population, reflecting the tension between the benefit of acquiring rare beneficial mutations and the cost of increased genetic load (Wielgoss et al., 2013; Swings et al., 2017). Nevertheless, even if the hypermutator phenotype is transient, FQ-R mutant *Mtb* strains with hypermutator phenotypes may be more adaptable than *Mtb* strains with normal mutation rates.

If hypermutator phenotypes are induced by FQ-R mutations in *Mtb*, their presence may be dependent on bacterial genetics. While the genetic diversity in *Mtb* is small compared with other bacterial pathogens, natural populations of *Mtb* can currently be grouped into seven distinct genetic lineages (Comas et al., 2010; Gagneux, 2018). This genetic diversity has been shown to modulate the frequency and phenotypes for INH-R (Gagneux et al., 2006a; Fenner et al., 2012; Ford et al., 2013), RIF-R (Gagneux et al., 2006c; Zaczek et al., 2009; Ford et al., 2013) and FQ-R (Castro et al., 2019). Therefore, the *Mtb* genetic background may modulate the frequency and magnitude of hypermutator phenotypes that would be caused by FQ-R mutations.

In this Chapter, we tested whether FQ-R mutations in DNA gyrase can induce a hypermutator phenotype in *Mtb*. Firstly, we used a fluctuation analysis as a measure of mutation rate, and tested whether FQ-R mutants would associate with higher frequencies of DR in vitro compared to their wild-type ancestor. We then used a collated set of publicly available genomic data from clinical isolates of *Mtb* to detect whether FQ-R mutations can modulate the frequency of mutations observed; specifically, we tested for an association between the presence of FQ-R mutations and increased genetic

diversity in natural populations of *Mtb*. Lastly, we tested whether FQ-R mutations can induce transient hypermutator phenotypes during the within-host evolution of *Mtb*. We used the genomic data from serially sampled *Mtb* isolates from MDR-TB patients from Georgia, and tested for an association between FQ-R mutations and changes in genetic diversity per serial *Mtb* isolate.

5.3 Methods

5.3.1 Collection of *Mtb* Strains for In Vitro Studies

We used the three genetically distinct, and fully drug-susceptible *Mtb* strains for our in vitro work: N0157 (a Lineage 1 or L1 strain), N1283 (Lineage 4 or L4), and N0145 (Lineage 2 or L2) (Comas et al., 2010; Gagneux, 2018). All strains were previously isolated from patients and characterized by Borrell et al., 2019 (Table 7.1, see Supplementary Information). We isolated spontaneous ofloxacin-resistant mutants from each of the three strains as described in the Methods of Chapter 3 and in Castro et al., 2019. The ofloxacin-resistant mutants harboured one of four ofloxacin-resistance *gyrA* mutations: G88C, A90V, D94G, or D94N. All strains were stored in -80 °C. Prior to experimentation, starter cultures were prepared for each *Mtb* strain by recovering a 20 μl aliquot from frozen stocks into a 10 ml volume of Middlebrook 7H9 broth (BD), supplemented with an albumin (Fraction V, Roche), dextrose (Sigma–Aldrich), catalase (Sigma–Aldrich), and 0.05% Tween 80 (AppliChem) (hereafter designated as 7H9 ADC). These starter cultures were incubated until their optical density at wavelength of 600 nm (OD_{600}) was $\sim$0.50, and were then used for in vitro assays.

5.3.2 Fluctuation Analyses

Fluctuation analysis experiments were designed based on Luria et al., 1943. Experiments and statistical analysis of data were performed as described as described in the Methods of Chapter 3 and in Castro et al., 2019, with the selective plates containing 100 μg/ml of streptomycin. Of particular note, significant differences between strains in their frequency of streptomycin-resistance mutations were defined by non-overlapping 95% confidence intervals (Rosche et al., 2000; Carey et al., 2018). Data analyses were performed using the R statistical software (v.3.5.1; R Core Team, 2018), and figures were produced using the ggplot2 package (Wickham, 2016).

5.3.3 Determining the Mutational Profile for Streptomycin-resistance In Vitro

From the parallel cultures plated on 100 μg/ml of streptomycin, the mutational profile for streptomycin-resistance was performed as described in the Methods of Chapter 3 and in Castro et al., 2019. Notably, heat-inactivated bacterial cultures were used as PCR templates to amplify the *rpsL* gene using the forward primer CGTGAAAGCGCCCAAGATAG and the reverse primer GAACCGCGGATGATCTTGTAG. This produces 333 bp PCR products, which were sent to Macrogen, Inc. or Microsynth AG for Sanger sequencing. Amino acid substitutions in the K43 or K88 codons of the

rpsL gene were identified by first aligning the PCR product against the H37Rv reference sequence using the Staden Package (Cole et al., 1998; Staden, 1996), and base-pair substitutions were identified using a custom R script (R Core Team, 2018). Fisher's exact test was used to test for significant differences between strains' mutational profiles for streptomycin-resistance. Figures were produced using the ggplot2 package (Wickham, 2016).

5.3.4 Whole Genome Sequencing Analysis of Publicly Available *M. tuberculosis* Genomic Sequences

Publicly available genomic sequences from clinical isolates of *Mtb* were collected as described in Menardo et al., 2018. Their accession numbers are listed in the Additional Files of Menardo et al., 2018, and not included here due to excessive length (n = 10,303). Publicly available genomic sequences from the study by Casali et al., 2014 were also downloaded and used for analysis.

All downloaded genomic sequences were re-analyzed using the sequence analysis pipeline described in Loiseau et al., 2019. Specifically, reads were trimmed with Trimmomatic (v.0.33; Bolger et al., 2014), and only reads longer than 20 bp were kept for analysis. Identification and merging of overlapping paired-end reads were performed using the SeqPrep software (https://github.com/jstjohn/SeqPrep). The MEM algorith of BWA (v.0.7.13; Li et al., 2009) was used to align the reads to the reconstructed ancestral sequence of *Mtb* (Comas et al., 2010). Duplicated reads were identified using the MarkDuplicates module of Picard (v.2.9.1; https://github.com/broadinstitute/picard). Local realignment of reads around Insertion/Deletions (INDELs) were performed using the RealignerTargetCreater and IndelRealigner modules of GATK (v.3.4.0; McKenna et al., 2010). Single nucleotide polymorphisms (SNPs) were called using the Samtools mpileup (v.1.2; Li, 2011) and VarScan (v.2.4.1; Koboldt et al., 2012) using the following thresholds: mapping quality ≥ 20, base quality at a given position ≥ 20, read depth at a given position $\geq 7\times$, and strand bias at a given position $\leq 90\%$. SNPs were defined as "fixed" in the population when it reached a frequency of $\geq$90%, and considered "variable" in the population when its frequency was between 10% and 90%. SNPs were annotated using snpEff (v.4.1144; Cingolani et al., 2012) using the *Mtb* H37Rv reference annotation (NC_000962.3).

5.3.5 Defining Transmission Clusters in Publicly Available *M. tuberculosis* Genomic Sequences

Transmission clusters were defined as described in the Methods of Chapter 3 and in Castro et al., 2019. Specifically, the haplotypes package (v.1.0) for the R statistical software (v.3.5.1; R Core Team, 2018) was used to infer a genetic distance matrix, using differences in the number of SNPs as a measure of genetic distance. Of particular note, a threshold of 5 SNPs average distance was used to define likely patient-to-patient transmission (Walker et al., 2013), and the tree was cut at a height of 5 SNPs using the hclust function. All resulting transmission clusters of strains, with a minimum size of three clustered genomes, were used for further analysis.

5.3.6 Measurement of Genetic Diversity

Genetic diversity as measured by pairwise nucleotide diversity (π) was performed as described by Nei et al., 1979 and Russell et al., 2017. Only fixed SNPs were used to measure π. The π for a given group of genomes was calculated per site and averaged across 10 kb non-overlapping windows, as in Equation (5):

$$\pi = \sum_{j=1}^{l} \frac{\frac{\sum_{i=1}^{a} x_i(n-x_i)}{n(n-1)}}{L} \tag{5}$$

Where n is the total number of genomes sampled and x_i is the number of genomes with allele i. The parameter a denoted the number of alleles present at locus j, and l denoted the total number of loci within a 10 kb window. The sequence length (10 kb) was denoted by L. Significant differences in the distribution of π per 10 kb non-overlapping regions between two groups of genomes were performed using Bonferonni-corrected Mann-Whitney U test.

Genetic diversity as measured by mean heterozygosity (H) was performed as described by Li, 1997 and Cuevas et al., 2015. Only variable SNPs were used to measure H. Per site heterozygosity (h_l) in a given genome was calculated according to Equation (6):

$$h_l = 1 - \sum_{i=1}^{a} f_{li}^2 \tag{6}$$

Where f_{li} is the frequency of allele i, and a is the total number of alleles present at locus l. The average heterozygosity of a given genome was calculated as in Equation (7):

$$H = \frac{1}{L} \sum_{l=1}^{L} h_l \tag{7}$$

Where L is equal to 4,411,529 bp, the length of the *Mtb* genome according to Cole et al., 1998. Significant differences in the distribution of H between two groups of genomes were performed using Bonferonni-corrected Mann-Whitney U test.

All measuments of genetic diversity were performed using custom R scripts (v.3.5.1; R Core Team, 2018), and figures were produced using the ggplot2 package (Wickham, 2016).

5.3.7 Calculating Terminal Branch Lengths in the Genomic Data Set from Casali et al., 2014

Terminal branch lengths for the genome data set from Casali et al., 2014 were calculated by first using RAxML (v.8.2.8; Stamatakis, 2014) to infer a maximum likelihood phylogeny using the variable SNP alignment file. We used the general time-reversible model of sequence evolution, and bootstrapped the highest scoring maximum likelihood tree (1000 pseudoreplicates) to infer branch support values. The phylogeny was rooted using the *M. canettii* sequence (accession number: SRR011186). Terminal branch length per genome was extracted from the resulting phylogenetic tree using a custom R script (v.3.5.1; R Core Team, 2018) that first identified the node numbers per terminal branch, and

then extracted the branch length that was associated with it.

5.3.8 Genomic Data of *M. tuberculosis* Isolates that were Serially-sampled from MDR-TB Patients from Georgia

Serial sputum samples were collected from 118 smear-positive MDR-TB patients from Georgia. All samples were part of the collection of the National Center for Tuberculosis and Lung Diseases, located in Tblisi, Georgia. Ethical review and approval for the use and analysis of these samples was part of the 309540-EVODRTB research agreement granted by the European Research Council. The given 118 MDR-TB patients were chosen for sample collection as these patients had slow sputum smear conversion, resulting in repeated visits. No patient data are currently available for these samples, with the exception of drug-susceptibility testing (DST) data, which was performed at variable intervals during patient treatment. At least 3 serial samples for a given patient had to be present to be included for downstream analysis. Sputum samples were decontaminated, and then grown on Löwenstein-Jensen (L-J) solid medium. Resulting colonies on the L-J medium were scraped, and DNA was extracted according to the CTAB method (Soolingen et al., 1991). DNA sequencing was performed on Illumina HiSeq 2000/2500 paired end technology (PRJNA488343), and only genomes with a minimum sequencing depth of 15× were kept. This provided a data set with 425 genomes, with a resulting median coverage per genome of 75× and an interquantile range of 40× to 110×. Genomes were grouped depending on whether the MDR-TB patients they were isolated from acquired additional drug-resistance during treatment based on DST results (DR Gain group), or did not gain additional resistance based on DST results (Control group).

5.4 Results

5.4.1 Fluoroquinolone-resistant *gyrA* mutations can increase frequencies of acquiring further streptomycin-resistance acquisition in *M. tuberculosis*

We tested whether FQ-R *gyrA* mutations can modulate the in vitro mutation rates of *Mtb*. To do so, we used the Luria-Delbrück fluctuation analysis to measure the frequency of resistance to a high concentration of streptomycin (STR; 100 μg/mL). We reasoned that this high STR concentration provided an accurate measurement of the DNA base-pair mutation rate in *Mtb* when using a fluctuation analysis model because of two features. Firstly, STR-resistance (STR-R) in *Mtb* is generally restricted to chromosomal base-pair mutations (Finken et al., 1993; Sreevatsan et al., 1996; Nhu et al., 2012; Jagielski et al., 2014). Secondly, the high STR concentration would further restrict the STR-R mutations to a handful of codon positions that can provide the necessary resistance level to grow in the presence of the drug, namely the codons K43 and K88 from the *rpsL* gene (Sreevatsan et al., 1996; Nhu et al., 2012; Sun et al., 2018). We previously performed the high-level STR-R fluctuation analysis on three genetically distinct and drug-susceptible *Mtb* strains: N0157 (Lineage 1 or L1, Manila sublineage),

N1283 (Lineage 4 or L4, Ural sublineage), and N0145 (Lineage 2 or L2, Beijing sublineage) (Castro et al., 2019). Using two biological replicates per *Mtb* strain, we observed no evidence for a difference in the frequency of high-level STR-R in N0157, N1283, or N0145 (Figure 3.3, see Chapter 3) (Castro et al., 2019). This suggested that there were minimal, if any, differences in their mutation rates.

We then isolated FQ-R mutants carrying one of four possible *gyrA* mutations (G88C, A90V, D94G, or D94N) in either N0157, N1283, or N0145. The frequency of high-level STR-R was measured for each GyrA mutant, with the number of biological replicates performed per GyrA mutant summarized in Table 5.1. We observed that the frequency of high-level STR-R for GyrA mutants was dependent on both the *gyrA* mutation present and the *Mtb* strain the mutation was present in (Figure 5.1). Firstly, the $GyrA^{D94G}$ mutants derived from two strains (N0157 and N0145) had higher frequencies of STR-R compared to their respective wild-type ancestral strain. Unfortunately, no fluctuation analysis data could be made available for N1283-derived $GyrA^{D94G}$ mutants due to potential contamination of the non-selective plates. In contrast to the $GyrA^{D94G}$ results, the $GyrA^{A90V}$ mutants derived from two strains (N1283 and N0145) showed no difference in their frequency of STR-R compared to their respective wild-type strains. All N1283-derived GyrA mutants acquired the same frequency of STR-R as their wild-type ancestor. Lastly, the $GyrA^{G88C}$ and $GyrA^{D94N}$ mutants derived from two strains (N0157 and N0145) produced contradicting results between their respective biological replicates (if more than one replicate was performed), with one replicate having a higher frequency of STR-R than their respective wild-type ancestor, while the other replicate showed no difference.

Table 5.1: Number of biological replicates performed per GyrA mutant per *M. tuberculosis* genetic background for the fluctuation analysis at 100 μg/mL streptomycin

Strain	Lineage	*gyrA* Mutation	Genetic Background (Ancestral Strain)	Number of replicates
N0157	L1	WT	—	2
N3661	L1	G88C	*N0157*	1
N2034	L1	A90V	*N0157*	0
N2036	L1	D94G	*N0157*	2
N2035	L1	D94N	*N0157*	2
N1283	L4	WT	—	2
N2508	L4	G88C	*N1283*	1
N2505	L4	A90V	*N1283*	1
N3915	L4	D94G	*N1283*	0
N2507	L4	D94N	*N1283*	1
N0145	L2	WT	—	2
N3659	L2	G88C	*N0145*	2
N2847	L2	A90V	*N0145*	1
N1893	L2	D94G	*N0145*	2
N1895	L2	D94N	*N0145*	2

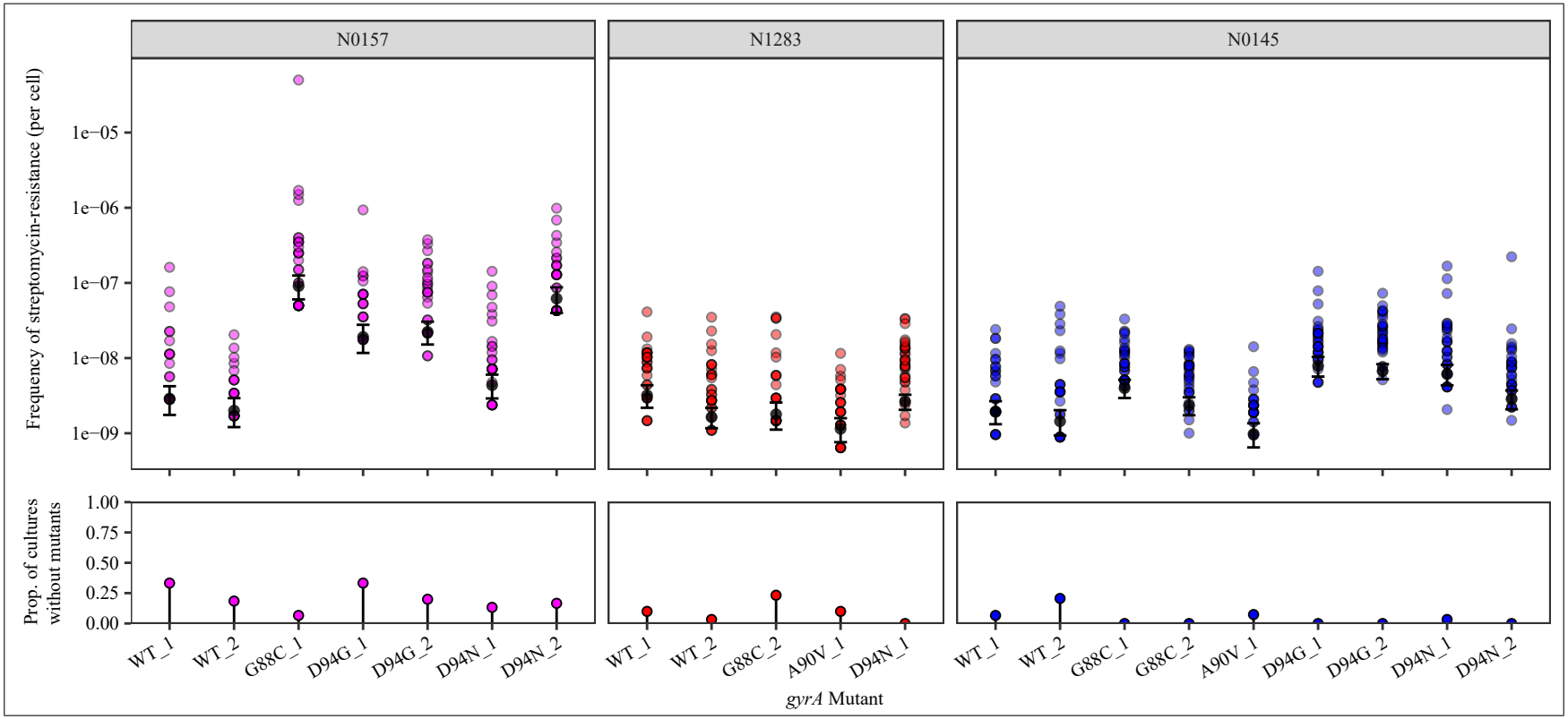


Figure 5.1: The frequency of streptomycin-resistance at 100 μg/mL for *M. tuberculosis* may be modulated by the presence of *gyrA* mutations. Top panel: coloured points represent the frequency of streptomycin-resistance per cell per parallel culture, with darker points representing multiple cultures with the same frequency. Colours denote the lineage that the *M. tuberculosis* strain belongs to (L1 = pink; L2 = blue; L4 = red). Grey points represent the estimated number of drug-resistance mutations per cell per strain as calculated by MSS-MLE, while black bars denote the respective 95% confidence intervals. Bottom panel: the percentage of parallel cultures lacking streptomycin-resistant mutants. Two biological replicates are presented for each *M. tuberculosis* strain, with each replicate identifier suffixed after the strain name.

We next explored the mutational profile for STR-R amongst the GyrA mutants and their respective wild-type ancestors by screening 760 resistant colonies for STR-R mutations. Specifically, we screened for mutations present in either the K43 or the K88 codons of the *rpsL* gene, as mutations in these codons confer high-level STR-R (Sreevatsan et al., 1996; Nhu et al., 2012; Sun et al., 2018). While no fluctuation analysis data were available for N1283-derived GyrA^{D94G} mutants due to potential contamination of the non-selective plates, the mutational profile for one replicate of the N1283-derived GyrA^{D94G} mutant was reported here (replicate "D94G_2"). We reasoned that any contamination would not survive the high STR concentration on the selective plates. Indeed, the *rpsL* sequences from the STR-R colonies from the N1283 replicate D94G_2 did not differ from the H37Rv reference sequence, with the exception of known STR-R mutations, if present, and were thus kept for analysis. We observed that the mutational profiles between the wild-type ancestral strains (N0157, N1283, and N0145) differed from each other (Figure 5.2; Table 5.2). For example, the majority of mutations observed in the N0157 strain was *rpsL* K43T, while the K43R mutation was the most prevalent in the N1283 and N0145 strains. For the GyrA mutants, we observed that STR-R mutational profiles were dependent on both the *gyrA* mutation present and in which *Mtb* strain that the *gyrA* mutation was present. Specifically, N0157-derived GyrA mutants had similar STR-R mutational profiles compared to their ancestor, with the *rpsL* K43T mutation being most common. The N0145-derived GyrA mutants also had similar STR-R mutational profiles as their ancestor, with the *rpsL* K43R mutation being most prevalent. In contrast, the N1283-derived GyrA mutants had different mutational profiles compared to their ancestor; while the *rpsL* K43R mutation made up the vast majority of mutations observed for wild-type N1283, the majority of STR-R colonies isolated from the GyrA mutants lacked any mutations in codons K43 and K88.

Therefore, our results show that FQ-R *gyrA* mutations can modulate the frequency of acquiring further STR-R, as well as the mutational profile for STR-R in vitro. We next tested whether these in vitro findings have relevance in a clinical setting.

5.4.2 Testing the impact of fluoroquinolone-resistance mutations on the genetic diversity present in natural populations of *M. tuberculosis*

In general, STR is no longer in use against *Mtb* infections (WHO, 2019a). Therefore, there are limited *Mtb* genomic data available that are both FQ-R and STR-R. Thus, we focused our investigation on the clinical relevance of our in vitro findings by testing whether the FQ-R mutations had an impact on *Mtb* mutation rates. Compared to bacterial strains with normal mutation rates, hypermutator strains are hypothesized to have higher levels of genetic diversity due to the higher mutation supply available (Taddei et al., 1997; Giraud et al., 2001; Wielgoss et al., 2013; Hughes et al., 2017). Therefore, we tested whether FQ-R mutations associated with changes in the genetic diversity in natural populations of *Mtb*. To do so, we collated a set of publicly available genomic sequences from clinical isolates of *Mtb*. We limited our analysis to L2 and L4 strains, as these two lineages had the highest

Table 5.2: Mutations present in the K43 and K88 codons of the *rpsL* gene for 760 streptomycin-resistant colonies following fluctuation analysis on 100 μg/mL of streptomycin (part 1 of 2).

Genetic Background (Ancestral Strain)	Lineage	GyrA Mutant Replicate	*rpsL* Mutation	Frequency
N0157	L1	WT_1	K43N	1
N0157	L1	WT_1	K43R	5
N0157	L1	WT_1	K43T	9
N0157	L1	WT_1	nc_nc	1
N0157	L1	WT_1	wt_wt	2
N0157	L1	G88C_1	K43M	2
N0157	L1	G88C_1	K43N	14
N0157	L1	G88C_1	K43R	19
N0157	L1	G88C_1	K43T	38
N0157	L1	G88C_1	wt_wt	3
N0157	L1	G88C_1	nc_wt	1
N0157	L1	D94N_2	K43R	14
N0157	L1	D94N_2	K43T	40
N0157	L1	D94N_2	wt_wt	1
N1283	L4	WT_2	K43M	4
N1283	L4	WT_2	K43N	4
N1283	L4	WT_2	K43R	73
N1283	L4	WT_2	K43T	13
N1283	L4	WT_2	K88E	3
N1283	L4	WT_2	K88R	5
N1283	L4	WT_2	wt_wt	9
N1283	L4	WT_2	nc_nc	2
N1283	L4	WT_2	wt_nc	1
N1283	L4	D94G_2	K43N	3
N1283	L4	D94G_2	K43R	1
N1283	L4	D94G_2	wt_wt	40
N1283	L4	D94G_2	nc_nc	3
N1283	L4	D94G_2	nc_wt	1
N1283	L4	D94N_1	wt_wt	44
N1283	L4	D94N_1	nc_wt	1
N1283	L4	D94N_1	wt_nc	1

Note: wt_wt, wild-type in both K43 an K88; nc_nc, sequence not conclusive in codon K43 nor K88; nc_wt, sequence not conclusive in K43, but wild-type in K88; wt_nc, wild-type in K43, but sequence not conclusive in K88

Table 5.2: Mutations present in the K43 and K88 codons of the *rpsL* gene for 760 streptomycin-resistant colonies following fluctuation analysis on 100 μg/mL of streptomycin (part 2 of 2).

Genetic Background (Ancestral Strain)	Lineage	GyrA Mutant Replicate	*rpsL* Mutation	Frequency
N0145	L2	WT_1	K43R	1
N0145	L2	WT_1	K43T	1
N0145	L2	WT_2	K43R	42
N0145	L2	WT_2	K43T	12
N0145	L2	WT_2	wt_wt	8
N0145	L2	WT_2	nc_nc	1
N0145	L2	WT_2	nc_wt	1
N0145	L2	G88C_1	K43M	2
N0145	L2	G88C_1	K43N	2
N0145	L2	G88C_1	K43R	40
N0145	L2	G88C_1	K43T	37
N0145	L2	G88C_1	wt_wt	1
N0145	L2	G88C_1	nc_nc	11
N0145	L2	D94N_1	K43M	1
N0145	L2	D94N_1	K43N	2
N0145	L2	D94N_1	K43R	63
N0145	L2	D94N_1	K43T	16
N0145	L2	D94N_1	wt_wt	1
N0145	L2	D94N_1	nc_nc	3
N0145	L2	D94N_1	nc_wt	10
N0145	L2	D94N_2	K43M	5
N0145	L2	D94N_2	K43N	1
N0145	L2	D94N_2	K43R	67
N0145	L2	D94N_2	K43T	58
N0145	L2	D94N_2	wt_wt	1
N0145	L2	D94N_2	nc_nc	10
N0145	L2	D94N_2	nc_wt	5

Note: wt_wt, wild-type in both K43 and K88; nc_nc, sequence not conclusive in codon K43 nor K88; nc_wt, sequence not conclusive in K43, but wild-type in K88; wt_nc, wild-type in K43, but sequence not conclusive in K88

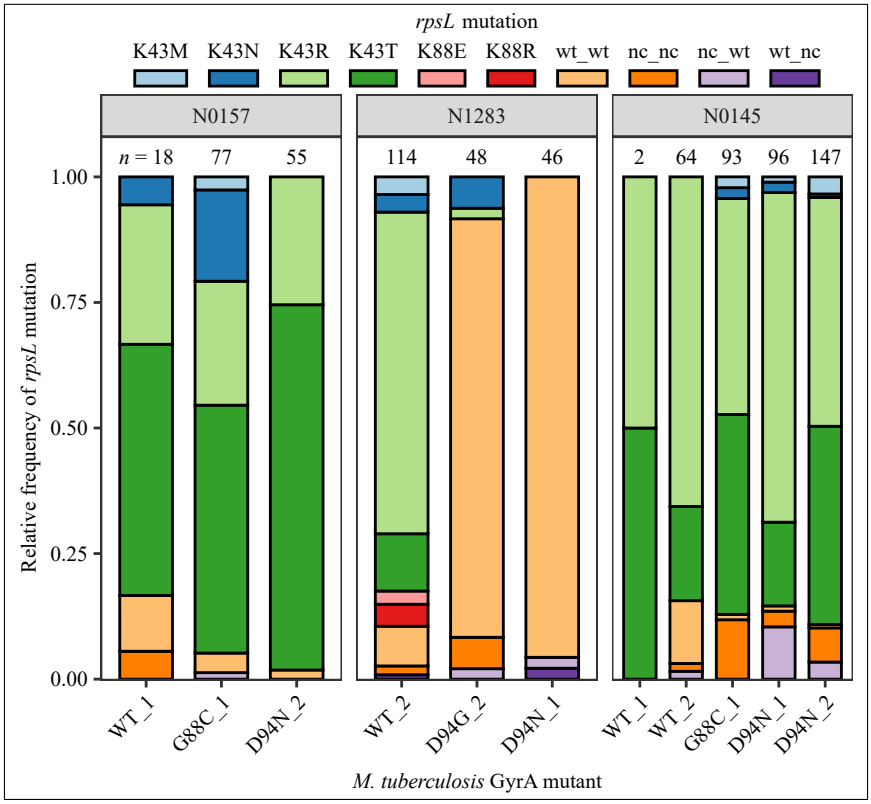


Figure 5.2: Variation in the mutational profile for streptomycin-resistance after fluctuation analysis on 100 μg/mL of streptomycin.
Mutations in the K43 and K88 codons of the *rpsL* gene was analyzed in 760 streptomycin-resistant colonies from the fluctuation analysis performed in Figure 5.1 (wt_wt, wild-type in both K43 an K88; nc_nc, sequence not conclusive in codon K43 nor K88; nc_wt, sequence not conclusive in K43, but wild-type in K88; wt_nc, wild-type in K43, but sequence not conclusive in K88). Mutational profiles for streptomycin-resistance of each GyrA mutant and their respective wild-type ancestors are shown here. The ancestral strain of each GyrA mutant is indicated in the grey bar above each panel. The biological replicates are presented for each GyrA mutant *M. tuberculosis* strain, with each replicate identifier suffixed after the *gyrA* mutation present in the strain. Numbers of colonies analyzed per *M. tuberculosis* genetic background are reported directly above each column.

clinical frequencies of FQ-R (Maruri et al., 2012; Castro et al., 2019). Due to standardized treatment regimens, the prevalence of FQ-R is heavily biased towards MDR-TB strains. Therefore, we analyzed only genomes that were classified as MDR-TB based on the presence of known RIF-resistance (RIF-R) *rpoB* mutations. We reasoned that RIF-R *rpoB* mutations are an acceptable genetic marker for MDR-TB, as they are generally acquired after INH-R in the clinic (Manson et al., 2017). Furthermore, we controlled for the potential bias that shared mutations in epidemiologically-linked *Mtb* strains would lead to lower observed genetic diversities (that is independent of FQ-R mutations) when compared to epidemiologically-independent *Mtb* strains. Therefore, we further filtered the data set to contain only genomes that were hypothesized to be epidemiologically-linked. We used differences in the number of single nucleotide polymorphisms (SNPs) as a measure of genetic distance between any two genomes,

Table 5.3: Number of clustered MDR-TB genomes per Lineage and per FQ-R group used in this analysis.

FQ-R Group	L2	L4
WT	1198	572
A90V	172	32
D94A	58	21
D94G	209	29
D94N	22	1
Other	151	37

and defined transmission clusters using a threshold of 5 SNPs average distance (Walker et al., 2013). This provided an initial set of clustered MDR-TB genomes with 1,810 genomes belonging to L2, and 692 genomes belonging to L4. The genomes were then assigned to a given group based on whether FQ-R mutations were absent (hereafter known as the "WT" group), or whether the fixed FQ-R *gyrA* mutations A90V, D94A, D94G or D94N were exclusively present. These mutation groups were chosen as they are four of the most prevalent FQ-R mutations in the clinic and in vitro (Maruri et al., 2012; Castro et al., 2019). These FQ-R groups provided reasonable sample sizes to test the association between FQ-R mutations and *Mtb* genetic diversity, with the exception of the D94N group in L4 where only 1 genome was present. Genomes that contained any other FQ-R mutations, including if they contained multiple FQ-R mutations, were simply assigned to the group "Other." The number of genomes per FQ-R group are presented in Table 5.3.

We measured the genetic diversity per FQ-R group (with the exception of the "Other" group, due to its heterogeneity in FQ-R mutations present). Specifically, we measured the pairwise nucleotide diversity (π) per FQ-R group, as described by Nei & Li (1979) and as performed by Russell & Cavanaugh (2017). For every FQ-R group, π was calculated per site and averaged across 10 kb nonoverlapping windows. We observed that for L2 strains, the distribution of π for the A90V, D94A, D94G, or D94N groups was significantly higher compared to their respective WT group (Figure 5.3, *P* for A90V = 0.03, *P* for D94A = 9.84×10^{-5}, *P* for D94G = 2.09×10^{-4}, *P* for D94N = 1.23×10^{-9}, Bonferonni-corrected Mann-Whitney *U* test). However, for L4 strains, the distribution of π was significantly lower in the A90V and D94A groups compared to their respective WT group (*P* for A90V = 1.40×10^{-19}, *P* for D94A = 1.06×10^{-59}, Bonferonni-corrected Mann-Whitney *U* test), while the D94G group showed no difference (*P* for D94G = 0.25, Bonferonni-corrected Mann-Whitney *U* test). No comparisons could be performed for the D94N group as only 1 genome was present. These genomic data results only partially supports our fluctuation analysis results as shown in Figure 5.1. Specifically, FQ-R *gyrA* mutations in L2 strains could increase mutation rates in vitro, and associated with increased genetic diversity in L2 strains in the clinic as measured by π. In contrast, we observed no evidence for FQ-R mutations modulating mutation rates in L4 strains in vitro, yet FQ-R mutations associated with decreased π in L4 strains in the clinic.

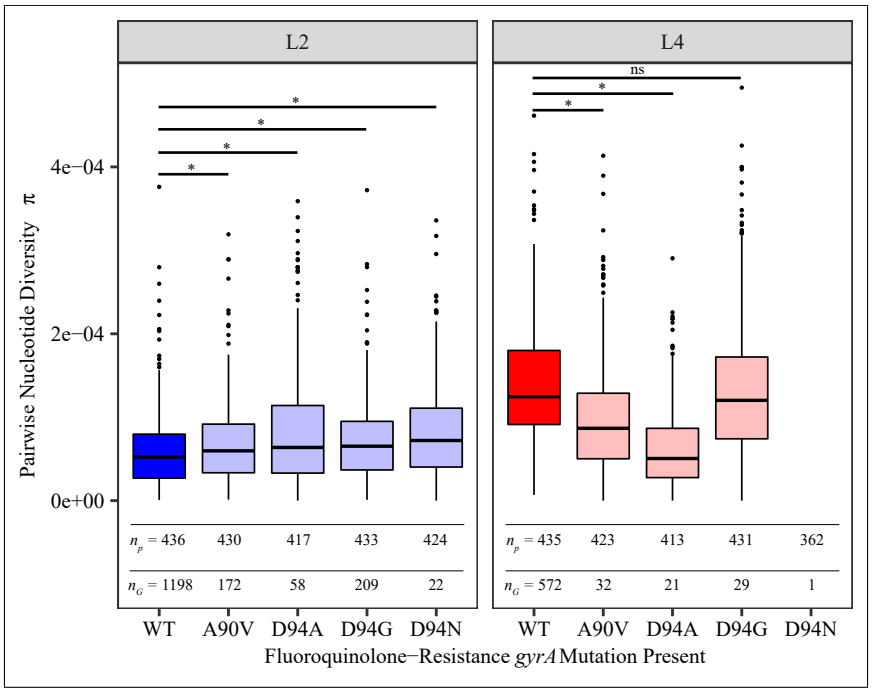


Figure 5.3: Variation in the genetic diversity amongst clustered *M. tuberculosis* strains isolated from MDR-TB patients, as measured by pairwise nucleotide diversity π.
Genomic data from clustered *M. tuberculosis* strains (cut-off = 5 SNPs average distance) belonging to either the L2 (blue hues) or L4 (red hues) lineages with known rifampicin-resistance *rpoB* mutations were included for analysis. Genomes were grouped based on whether they lacked FQ-R mutations (WT; darker hues), or if one of the following FQ-R *gyrA* mutations were exclusively present: A90V, D94A, D94G, or D94N (lighter hues). Values of π per genome group were calculated per 10 kb nonoverlapping window along the reconstructed ancestral *M. tuberculosis* genome, as described in Nei et al., 1979; Russell et al., 2017. The distribution of π values were illustrated using boxplots. Each box corresponds to the 25% and 75% quantiles, the black line represents the median, and the whiskers extend to 1.5 times the interquartile range. Number of 10 kb non-overlapping windows that π could be calculated for per group (denoted as n_P), as well as the number of genomes analyzed per group (denoted as n_G) were reported below each respective box plot. Significance was tested using Bonferroni-corrected pairwise Mann-Whitney *U* test.

The genetic diversity measured at the population-level, as is the case for using π, is not only a function of the individual strain mutation rates, but of population structure as well. Although our current data set includes only clustered genomes, this clustering is based on an average SNP distance of 5. Any two genomes may have larger distances between them, so long as the average distance between a given group of genomes is less than 5 SNPs. Therefore, even if there were no inherent differences in mutation rates, large distances between individual genomes in a set of clustered genomes would lead to higher observed values for π, and vice-versa. To control for this bias, we also measured the genetic diversity of each FQ-R group at the per-genome level. Specifically, we measured the mean heterozygosity (H) per genome in each FQ-R group, as described by Li (1997). We observed no evidence for a difference in the distribution of H between the A90V, D94A, D94G, or D94N groups compared to the WT group in L2 (Figure 5.4) (P for A90V = 1.72, P for D94A = 0.72, P for D94G = 1.27, P for D94N = 0.20, Bonferonni-corrected Mann-Whitney U test). We also observed no evidence for a difference in the distribution of H between the A90V, D94A, and D94G group compared to the WT group in L4 (P for A90V = 0.25, P for D94A = 0.25, P for D94G = 0.37, Bonferonni-corrected Mann-Whitney U test). While no tests could be performed to test for significance between the D94N group and the WT group for L4, the single measure of H for the D94N group was well within the 25% and 75% interquantile range of the WT group. These results suggest that FQ-R mutations do not have an impact on the genetic diversity of these clustered MDR-TB strains when using per-genome level measures. Therefore, FQ-R mutations do not likely modulate mutation rates in these clustered MDR-TB genomes.

Although the clustered MDR-TB genomes are closely related genetically, they were collected from a global sample set. We next aimed to analyze a genomic data set that was collected from a specific regional area. We chose the data set collected by Casali et al. (hereafter designated as "Casali 2014 data set"), as it provided a large collection of genomes that were prospectively collected between 2008 to 2010 in Samara Oblast, Russia (Casali et al., 2014). We initially analyzed 1,079 genomes from the Casali 2014 data set (Table 5.4). However, due to the few numbers of genomes with FQ-R mutations in L4 strains, we focused our analysis on L2 strains, and again grouped the genomes according the presence of specific FQ-R mutations. We measured their per-genome genetic diversity via mean heterozygosity (H) (Li, 1997), and found no evidence for a difference in the distribution of H when comparing the A90V, D94A, D94G, or D94N groups to the WT group in L2 strains (Figure 5.5; P for A90V = 3.52, P for D94A = 2.68, P for D94G = 0.66, P for D94N = 2.71, Bonferonni-corrected Mann-Whitney U test). These results suggest that FQ-R mutations do not have an impact on the genetic diversity of the *Mtb* strains from the Casali 2014 data set, and likely do not modulate mutation rates in these strains.

To further test whether FQ-R mutations have an impact on strain mutation rates from the Casali 2014 data set, we used differences in the terminal branch lengths of each strain in a phylogeny as another measure of strain-specific mutation rates. We hypothesized that for a given phylogenetic tree, *Mtb* strains with higher mutation rates would have longer terminal branch lengths due to the higher rate of mutation acquisition. Therefore, if FQ-R mutations increased mutation rates, then FQ-R mutations

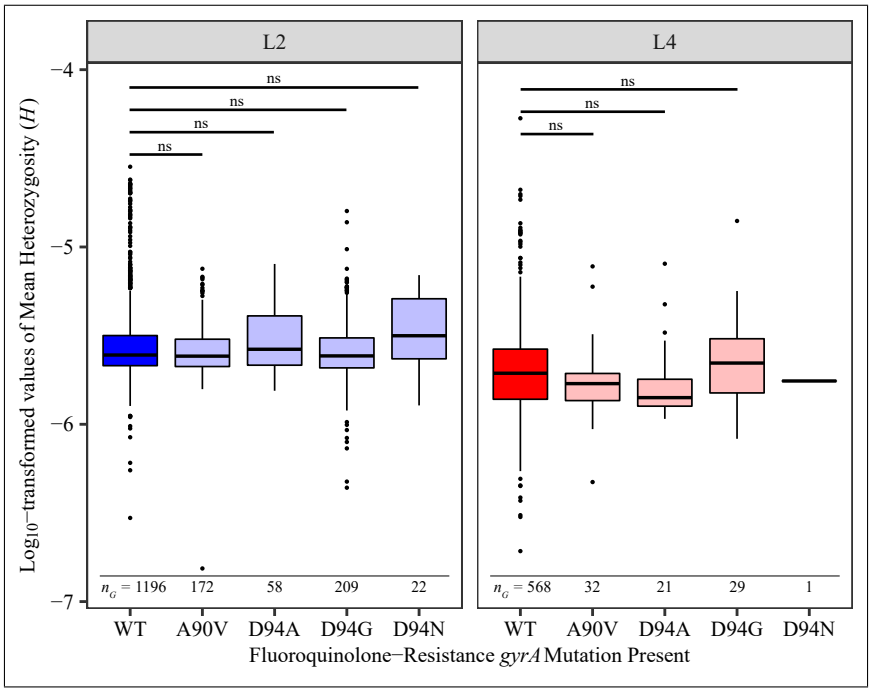


Figure 5.4: The genetic diversity amongst clustered *M. tuberculosis* strains isolated from MDR-TB patients, as measured by mean heterozygosity H.
Genomic data from clustered *M. tuberuclosis* strains (cut-off = 5 SNPs average distance) belonging to either the L2 (blue hues) or L4 (red hues) lineages with known rifampicin-resistance *rpoB* mutations were included for analysis. Genomes were grouped based on whether they lacked FQ-R mutations (WT; darker hues), or if one of the following FQ-R *gyrA* mutations were exclusively present: A90V, D94A, D94G, or D94N (lighter hues). The H per genome were calculated as described in Li, 1997; Cuevas et al., 2015. The distribution of H values were illustrated using boxplots. Each box corresponds to the 25% and 75% quantiles, the black line represents the median, and the whiskers extend to 1.5 times the interquartile range. Number of genomes analyzed per group (denoted as n_G) were reported below each respective box plot. Significance was tested using Bonferroni-corrected pairwise Mann-Whitney U test.

Table 5.4: Number of genomes per Lineage and FQ-R grouping in the Casali *et al.*, 2014 data set used in this analysis.

FQ-R Group	L2	L4
WT	587	368
A90V	11	2
D94A	13	1
D94G	38	1
D94N	7	0
Other	49	2

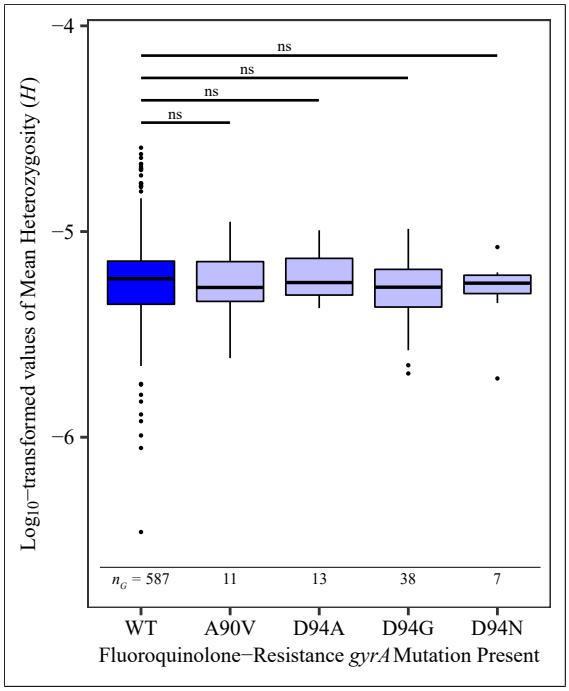


Figure 5.5: The genetic diversity amongst clinical *M. tuberculosis* strains from the Casali 2014 data set, as measured by mean heterozygosity H.
Only genomic data from *M. tuberculosis* strains belonging to L2 were included for analysis. Genomes were grouped based on whether they lacked FQ-R mutations (WT; dark blue), or if one of the following FQ-R *gyrA* mutations were exclusively present: A90V, D94A, D94G, or D94N (light blue). The H per genome were calculated as described in Li, 1997. The distribution of H values were illustrated using boxplots. Each box corresponds to the 25% and 75% quantiles, the black line represents the median, and the whiskers extend to 1.5 times the interquartile range. Number of genomes analyzed per group (denoted as n_G) were reported below each respective box plot. Significance was tested using Bonferroni-corrected pairwise Mann-Whitney U test.

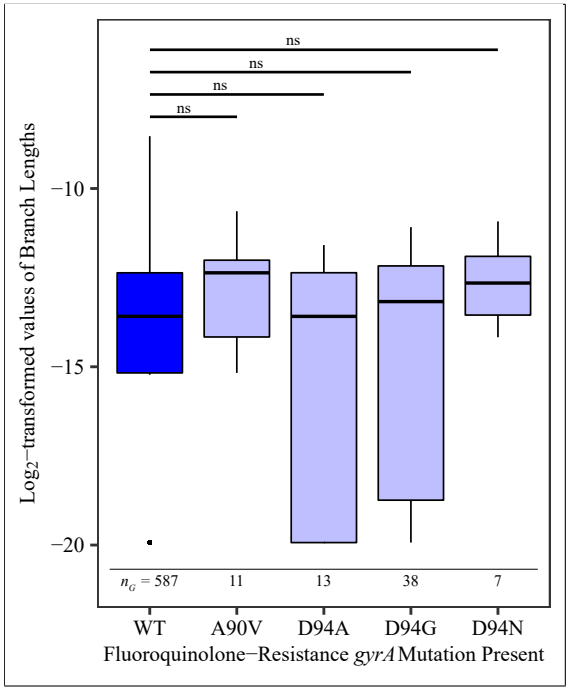


Figure 5.6: Edge length of terminal branches amongst clinical *M. tuberculosis* strains from the Casali 2014 data set.
Only genomic data from *M. tuberuclosis* strains belonging to L2 were included for analysis. Genomes were grouped based on whether they lacked FQ-R mutations (WT; dark blue), or if one of the following FQ-R *gyrA* mutations were exclusively present: A90V, D94A, D94G, or D94N (light blue). The distribution of branch lengths were illustrated using boxplots. Each box corresponds to the 25% and 75% quantiles, the black line represents the median, and the whiskers extend to 1.5 times the interquartile range. Number of genomes analyzed per group (denoted as n_G) were reported below each respective box plot. Significance was tested using Bonferroni-corrected pairwise Mann-Whitney *U* test.

would associate with longer terminal branch lengths. To test this hypothesis, we used the variable sites information from mapping the Casali 2014 data set to a reconstructed ancestral genome to construct a maximum-likelihood phylogeny, and calculated the terminal branch lengths for each strain. We found no evidence for a difference in terminal branch lengths between the A90V, D94A, D94G, or D94N groups compared to the WT group in L2 (Figure 5.6; *P* for A90V = 0.39, *P* for D94A = 2.92, *P* for D94G = 3.78, *P* for D94N = 0.61, Bonferonni-corrected Mann-Whitney *U* test). This further supports our conclusion that FQ-R mutations do not have an impact on the *Mtb* mutation rates from the Casali 2014 data set.

Hypermutator phenotypes may be transient in a population and gradually lost after the majority of mutations with large adaptive potential are acquired (Taddei et al., 1997; Desai et al., 2011; Lynch et al., 2016; Wielgoss et al., 2013; Swings et al., 2017). Any transient mutation rate changes would not be captured using the clustered MDR-TB or Casali 2014 data sets, as they have limited temporal information with each *Mtb* strain sampled for genomic information using only one time-point. Thus, we tested for an association between FQ-R mutations and transient hypermutator phenotypes by using

Table 5.5: Number of patient cases included per DST Group and FQ-R Index, and number of genomes analyzed per patient from the Georgian MDR-TB Serial Isolates data set

DST Group	FQ-R Index	Number of serially sampled MDR-TB patients included	Number of genomes analyzed per patient	Mean number of genomes analyzed per patient
Control	FQ-R: Initial	19	62	3.26
Control	FQ-R: No Gain	47	153	3.26
Control	FQ-R: Gained	2	8	4
DR Gain	FQ-R: Initial	8	26	3.25
DR Gain	FQ-R: No Gain	12	39	3.25
DR Gain	FQ-R: Gained	26	113	4.35

Mtb genomes that were serially sampled from 118 TB patients from the country of Georgia. These TB patients were all classified as MDR-TB based on drug-susceptibility testing (DST). Patients were serially sampled for *Mtb* isolates during treatment, with a total of 547 *Mtb* samples collected between 2009 and 2013. The TB patient cases were grouped into whether their DST results went unchanged during the treatment and serial sampling (DST Group: Control), or whether they gained resistance to kanamycin, ofloxacin, ethambutol, capreomycin, para-aminosalicylic acid, or any of these combinations (DST Group: DR Gain). We performed a preliminary analysis, as of the 547 samples collected, only 425 genomes with a sequencing coverage of at least 15× could be made presently available. The median coverage per isolate was 75×, with an interquantile range of 40× to 110×. The genomes were then grouped into whether FQ-R mutations were already present at the earliest available serial isolate (FQ-R Index: Initial), no FQ-R mutations were acquired during serial sampling (FQ-R Index: No Gain), or FQ-R mutations were acquired (FQ-R Index: Gained). We summarized the number of genomes analyzed per group in Table 5.5, and illustrated the sampling dates per serial isolate for the DR Gain group in Figure 5.7. Notably, two TB cases from the Control group (*i.e.* did not gain any further drug-resistance during treatment, as defined by DST) actually acquired and then subsequently lost FQ-R mutations.

We first calculated the mean heterozygosity (H) per genome from the *Mtb* serial isolates (Li, 1997). We observed some variability in H through the serial sampling for all DST groupings and FQ-R Indices (Figures 5.8 & 5.9). To detect transient hypermutator phenotypes, we calculated the differentials in H by substracting the H of a given genome sample by the H of the genome sample previous to it in its respective series. We illustrate the differential H per serial isolate in Figures 5.10 & 5.11. Transient hypermutator phenotypes due to FQ-R mutations may be identifiable by the presence of a positive association between FQ-R mutations and positive differential H. We found no evidence for a difference in differential H between genomes with or without FQ-R mutations, irrespective of the FQ-R Index grouping (Figure 5.12; $P = 0.46$ for the FQ-R Index: Initial group, $P = 0.25$ for the FQ-R Index: Gained group, Welch's Two Sample t-Test). Thus, FQ-R mutations did not associate with positive differential H. These results suggest that *Mtb* populations that acquired FQ-R mutations during treat-

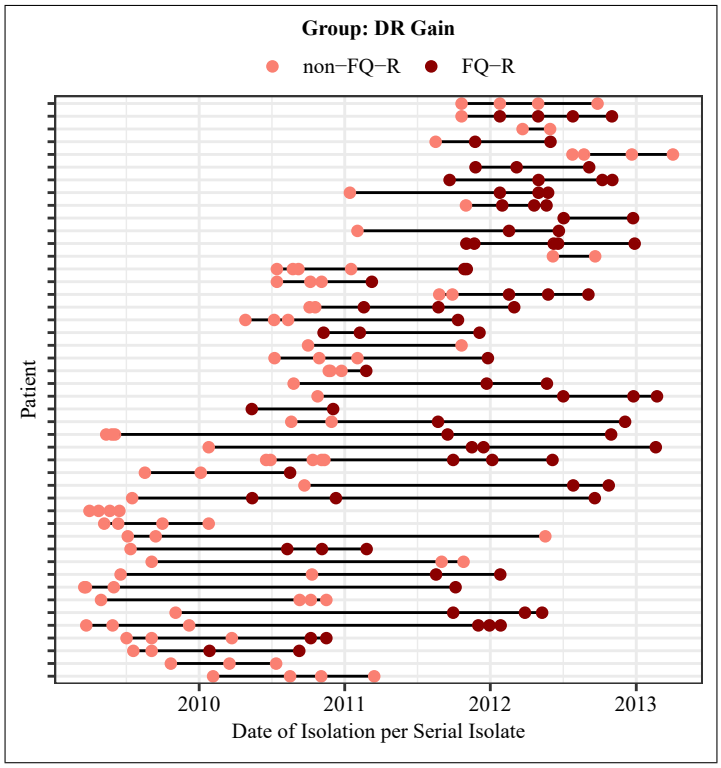


Figure 5.7: Isolation Dates for *M. tuberculosis* genomes that were serially sampled from MDR-TB patients from the country of Georgia.
Coloured points denote the isolation date of a given serial sample, with the colours denoting the presence (dark red) or absence (salmon) of FQ-R mutations. Serial isolates for a given patient are connected by black lines. Only the isolation dates for the *M. tuberculosis* genomes from the DR Gain group are represented here (*i.e.* Patient group whose DST results changed during treatment due to the acquisition of further drug-resistance to kanamycin, ofloxacin, ethambutol, capreomycin, para-aminosalicylic acid, or any of these combinations).

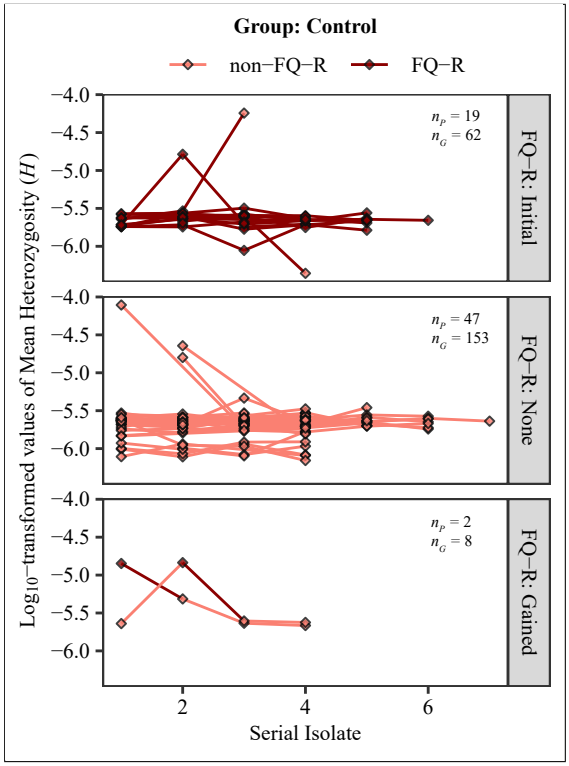


Figure 5.8: Mean heterozygosity (H) per serial *M. tuberculosis* isolate per MDR-TB patient from the Control group.
Patients were classified into the Control group if their DST results remained unchanged during treatment. Coloured points signify H of an *M. tuberculosis* isolate, while coloured lines connect serial isolates belonging to a given patient. Colours denote the presence (dark red) or absence (salmon) of FQ-R mutations in a given serial isolate. Serial isolates were further grouped into an FQ-R Index, which is based on whether the earliest serial isolate had a FQ-R mutation ("Initial"), had no FQ-R mutation ("None"), or gained FQ-R mutation at any point during the treatment, even if the FQ-R mutations were eventually lost ("Gained"). Number of patients (denoted as n_P) and number of genomes (denoted as n_G) analyzed per group were reported in top right of each panel.

ment were not any more likely to undergo transient hypermutator phenotypes than *Mtb* populations without FQ-R mutations.

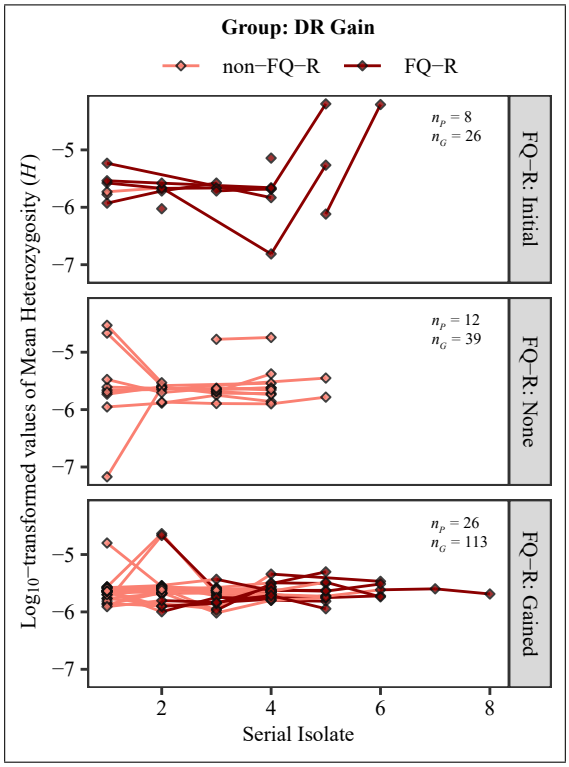


Figure 5.9: Mean heterozygosity (H) per serial *M. tuberculosis* isolate per MDR-TB patient from the DR Gain group.

Patients were classified into the DR Gain group if they acquired further DR phenotypes during treatment according to DST. Coloured points signify H of an *M. tuberculosis* isolate, while coloured lines connect serial isolates belonging to a given patient. Colours denote the presence (dark red) or absence (salmon) of FQ-R mutations in a given serial isolate. Serial isolates were further grouped into an FQ-R Index, which is based on whether the earliest serial isolate had a FQ-R mutation ("Initial"), had no FQ-R mutation ("None"), or gained FQ-R mutation at any point during the treatment, even if the FQ-R mutations were eventually lost ("Gained"). Number of patients (denoted as n_P) and number of genomes (denoted as n_G) analyzed per group were reported in top right of each panel.

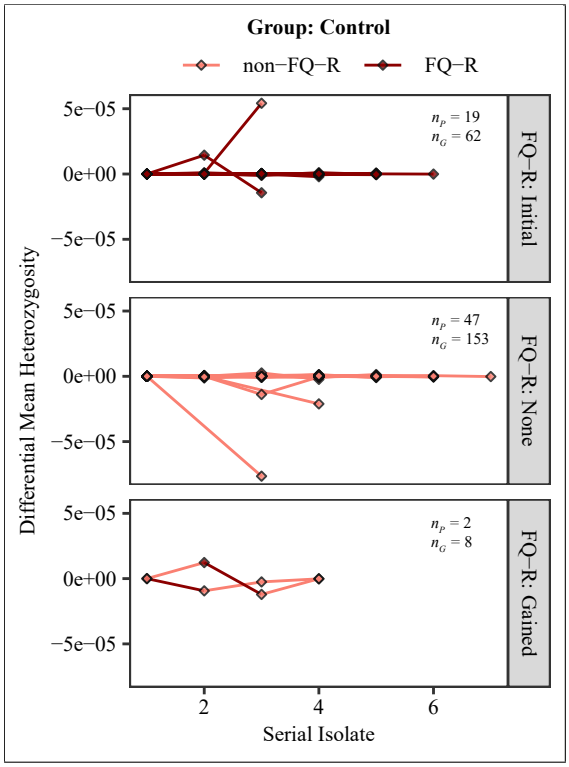


Figure 5.10: Differential mean heterozygosity (H) per serial *M. tuberculosis* isolate per MDR-TB patient from the Control group.
Patients were classified into the Control group if their DST results remained unchanged during treatment. Coloured points signify differential H of an *M. tuberculosis* isolate, while coloured lines connect serial isolates belonging to a given patient. Serial isolates are coloured based on the presence (dark red) or absence (salmon) of FQ-R mutations. Serial isolates were further grouped into an FQ-R Index, which is based on whether the earliest serial isolate had a FQ-R mutation ("Initial"), had no FQ-R mutation ("None"), or gained FQ-R mutation at any point during the treatment, even if the FQ-R mutations were eventually lost ("Gained"). Number of patients (denoted as n_P) and number of genomes (denoted as n_G) analyzed per group were reported in top right of each panel.

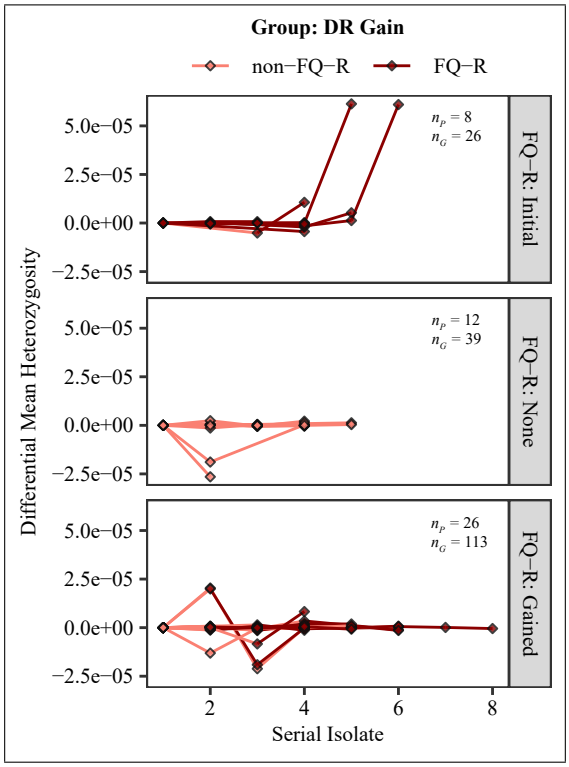


Figure 5.11: Differential mean heterozygosity (H) per serial *M. tuberculosis* isolate per MDR-TB patient from the DR Gain group.
Patients were classified into the DR Gain group if they acquired further DR phenotypes during treatment according to DST. Coloured points signify differential H of an *M. tuberculosis* isolate, while coloured lines connect serial isolates belonging to a given patient. Serial isolates are coloured based on the presence (dark red) or absence (salmon) of FQ-R mutations. Serial isolates were further grouped into an FQ-R Index, which is based on whether the earliest serial isolate had a FQ-R mutation ("Initial"), had no FQ-R mutation ("None"), or gained FQ-R mutation at any point during the treatment, even if the FQ-R mutations were eventually lost ("Gained"). Number of patients (denoted as n_P) and number of genomes (denoted as n_G) analyzed per group were reported in top right of each panel.

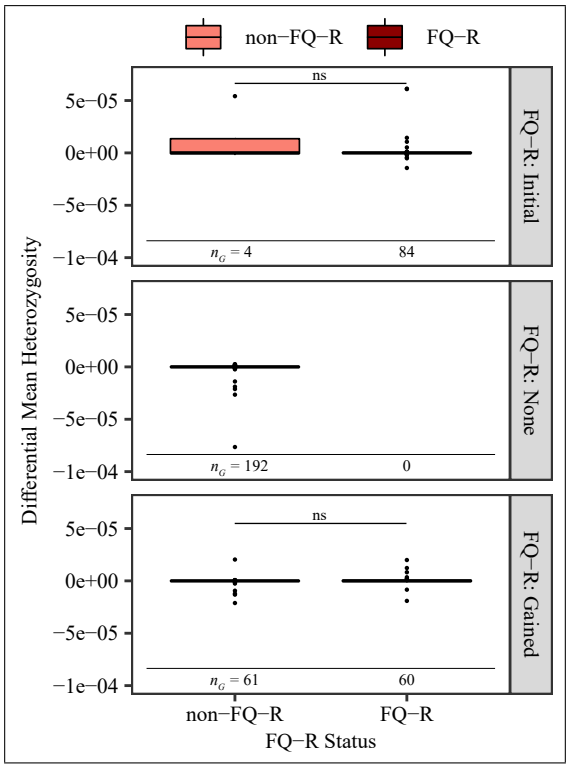


Figure 5.12: Distribution of differential mean heterozygosities (differential H) per *M. tuberculosis* serial isolate versus the presence or absence of fluoroquinolone-resistance mutations.
Differential H per *M. tuberculosis* serial isolate were first grouped by FQ-R Index, which is based on whether the earliest serial isolate had a FQ-R mutation ("Initial"), had no FQ-R mutation ("None"), or gained FQ-R mutation at any point during the treatment, even if the FQ-R mutations were eventually lost ("Gained"). Differential H were further grouped and coloured based on the presence (dark-red) or absence (salmon) of FQ-R mutations. The distribution of differential H values were illustrated using boxplots. Each box corresponds to the 25% and 75% quantiles, the black line represents the median, and the whiskers extend to 1.5 times the interquartile range. Significance was tested using Welch's Two Sample t-Test (alpha = 0.05). Number of genomes analyzed per group (denoted as n_G) were reported below each respective box plot.

5.5 Discussion

In this Chapter, we showed that specific FQ-R mutations can modulate the rate of acquiring further drug-resistance mutations in *Mtb* in vitro. However, analysis of publicly available genomic data from clinical isolates of *Mtb* showed no evidence for an association between FQ-R mutations and increased genetic diversity. Further, analysis of genomic data from isolates that were serially sampled from MDR-TB patients from Georgia showed no evidence for FQ-R mutations associating with transient hypermutator phenotypes during treatment. Thus, our results suggest that FQ-R mutations do not generally induce hypermutator phenotypes, nor lead to higher genetic diversity in *Mtb*.

We first used a fluctuation analysis on high-level STR-R as a proxy measurement for detecting hypermutator phenotypes in three genetically-distinct wild-type *Mtb* strains (N0157, N1283, and N0145) and in their respective FQ-R mutants with one of four possible *gyrA* mutations (G88C, A90V, D94G, or D94N). Our results showed that, with the exception of the *gyrA* D94G mutation, FQ-R mutations generally conferred no mutation rates differences (Figure 5.1). Previous work attempting to associate FQ-R mutations with hypermutator phenotypes in other bacteria have provided similar results. Specifically, Gould *et al.* have shown that clinical isolates of FQ-R *Escherichia coli* may have hypermutator phenotypes (Gould et al., 2007). However, this hypermutator phenotype was associated with missense mutations in the DNA repair gene *mutS* (Gould et al., 2007), a gene where mutations are known to confer a hypermutator phenotype (Chopra et al., 2003; Maciá et al., 2005; Oliver et al., 2010). In contrast, the most common FQ-R mutations in *gyrA*, *parC*, and *parE* did not associate with increased mutation rates in *E. coli*. Therefore, while specific FQ-R *gyrA* mutations may modulate mutation rates, FQ-R *gyrA* mutations do not generally induce hypermutator phenotypes.

However, our results show that the presence of a hypermutator phenotype was dependent on both the FQ-R *gyrA* mutation present. Indeed, only the FQ-R *gyrA* D94G mutation clearly associated with increased frequencies of STR-R, irrespective of in which strain it was present. This is in accordance with our previous work showing that FQ-R mutation phenotypes are dependent on the given FQ-R mutation present (Castro et al., 2019).

An epistatic interaction between the FQ-R *gyrA* mutation and *Mtb* genetic background had a notable effect on the mutational profile for STR-R. Firstly, *rpsL* mutations in codon K43 (most notably, K43R and K43T) made up the vast majority of STR-R mutations present after fluctuation analysis of the N0157 and N0145 strains, as well as in their respective GyrA mutants (Figure 5.2; Table 5.2). This was as expected, as mutations in the K43 and K88 codon of *rpsL* are known to confer high-level STR-R (Sreevatsan et al., 1996; Nhu et al., 2012). However, while the STR-R *rpsL* K43R mutation was most prevalent in the wild-type N1283 strain, the vast majority of STR-R colonies present in the N1283-derived GyrA mutants had no mutations in the *rpsL* K43 nor the K88 codons. This suggested that two types of epistasis modulate the evolution of STR-R in our in vitro system. First, there was an epistatic interaction between FQ-R *gyrA* and STR-R *rpsL* mutations, as the N1283-derived GyrA mutant strains had different mutational profiles for STR-R compared to their wild-type ancestor. This may

be due to *rpsL* mutations conferring different levels of resistance to STR when they were co-present with *gyrA* mutations. Epistatic interactions between chromosomal DR mutations have been reported in the literature. The phenotype of STR-R *rpsL* mutations differed when they were in the presence of RIF-R *rpoB* mutations in *E. coli* (Trindade et al., 2009; Angst et al., 2013; Durão et al., 2015). Of particular note, the phenotypes of STR-R *rpsL* mutations in *M. smegmatis* differed depending on which FQ-R *gyrA* mutations were present (Sun et al., 2018). This suggests that epistasis between STR-R *rpsL* mutations and FQ-R *gyrA* mutations may be conserved amongst mycobacteria.

The second epistatic interaction that governed the evolution of STR-R in our in vitro system was an interaction between chromosomal DR mutations and the bacterial genetic background. Specifically, only GyrA mutants derived from the N1283 strain differed in their STR-R mutational profiles compared to their respective wild-type ancestor. The role of the genetic background in modulating DR mutational profiles has been reported in the literature. The genetic background has been shown to modulate the mutational profile for INH-R (Gagneux et al., 2006a; Fenner et al., 2012) and for FQ-R (Castro et al., 2019) in *Mtb*. The genetic background also modulated the mutational profile for RIF-R in *Pseudomonas spp.* (Vogwill et al., 2014; Vogwill et al., 2016) and in *E. coli* (Angst et al., 2013). These previous studies, along with the results shown here, suggest that the epistasis between chromosomal DR mutations and the bacterial genetic background, as well as epistasis between different chromosomal DR mutations, are general phenomena in bacterial DR evolution.

We next explored the clinical relevance of our in vitro work, and focused on whether the presence of FQ-R mutations can affect mutation rates in natural populations of *Mtb* by using levels of genetic diversity as a proxy measure. We first used the average pairwise nucleotide diversity (π) as a measure of genetic diversity (Nei et al., 1979) in a collated set of genomes isolated from clustered MDR-TB strains belonging to either L2 or L4. We found that while L2 strains with FQ-R mutations had higher π than L2 strains without FQ-R mutations, L4 strains with FQ-R mutations did not associate with higher π than L4 strains without FQ-R mutations (Figure 5.3). Rather than inherent differences in mutation rates, we hypothesize that this is likely due to differences in the population structure between L2 and L4 strains that developed FQ-R. This is due to two, non-exclusive reasons. Firstly, L2 strains are more associated with MDR-TB than L4 strains (Borrell et al., 2009; Fenner et al., 2012; Casali et al., 2014; Merker et al., 2015). We thus hypothesize that L2 MDR-TB strains have been transmitting between patients for a longer time than L4 strains, leading to L2 strains becoming FQ-R earlier than L4 strains. This would consequently allow the FQ-R L2 strains more time to acquire subsequent mutations and further diverge from one another, leading to the higher π observed in FQ-R L2 strains compared to FQ-R L4 strains. A second reason that we hypothesize that the differences in nucleotide diversity may be due population structure rather than inherent mutation rate differences comes from when we use the mean heterozygosity (H) per *Mtb* strain as another measure of genetic diversity (Li, 1997). Specifically, we observed that for both L2 and L4 strains in the clustered MDR-TB genomes, there were no differences in the H of strains with FQ-R mutations compared to strains without FQ-

R mutations (Figure 5.4). We controlled for a geographic sampling bias in the clustered MDR-TB genomes by using only *Mtb* genomes from the Casali 2014 data set, and again observed no differences in H between strains with or without FQ-R mutations (Figure 5.5) (Casali et al., 2014). Thus, FQ-R mutations do not appear to increase genetic diversity, and consequently may not lead to a stable hypermutator phenotype in natural populations of *Mtb*.

However, hypermutator phenotypes may be transient in a population (Taddei et al., 1997; Desai et al., 2011; Wielgoss et al., 2013; Lynch et al., 2016; Swings et al., 2017), and transient hypermutator phenotypes may only be observed using serial samples with temporal information. Therefore, we tested for the presence of transient hypermutator phenotypes conferred by FQ-R mutations by using genomes from serially sampled *Mtb* isolates from MDR-TB patients that were undergoing treatment in Georgia. While using H as the measure of genetic diversity, we showed some variability in genetic diversity was present between serial isolates per patient (Figures 5.10, 5.11). However, there was no evidence for FQ-R mutations associating with positive value changes in H (Figure 5.12). Therefore, FQ-R mutations do not appear to lead to transient hypermutator phenotypes *Mtb* populations within patients undergoing treatment for MDR-TB.

Our study was limited by the fact that no fluctuation analysis data for the N1283-derived D94G mutant could be made available due to potential contamination in the non-selective solid media in separate experiments. Therefore, it is unclear whether the *gyrA* D94G mutation would consistently lead to increased STR-R frequencies across the *Mtb* genetic background we tested.

The timing between the sampling of the serial isolates from the MDR-TB patients may limit our ability to detect transient hypermutator phenotypes. Indeed, FQ-R samples may be separated from previously non-FQ-R samples by up to two years in some patient cases (Figure 5.7). It has been shown that hypermutator phenotypes can emerge and then be lost up to three times within 200 generations of an *E. coli* population evolving in the presence of environmental stress (Swings et al., 2017). While *Mtb* populations in vitro can have generation times ranging from approximately 16 to 26 hours (Castro et al., 2019), their exact generation times within patients are currently unknown due to the extremely invasive procedure such a measurement would require. If the *Mtb* populations within the MDR-TB patients underwent approximately 200 generations between the sampling times, then any transient hypermutator phenotypes would not be captured due to the lack of temporal resolution. Therefore, a greater temporal resolution may be required to confidently conclude whether FQ-R mutations can induce a transient hypermutator phenotype.

Lastly, we focused our detection of hypermutator phenotypes using differences in DNA base-pair mutation rates. However, DNA base-pair mutations are not the only consequence of error-prone DNA repair mechanisms. For example, if the predominant DNA damage produced by FQ-R *gyrA* mutations are DSBs, then the hypermutator phenotypes would primarily manifest in increased insertion/deletion (INDELs) mutations. This is because DNA damage in the form of DSBs are repaired through one of the three DNA repair pathways in *Mtb*: homologous recombination (HR), non-

homologous end joining (NHEJ), or single-strand annealing (SSA) (Gupta et al., 2011; Singh, 2017). Of these repair mechanisms, NHEJ and SSA are the most error-prone (Gong et al., 2005; Gupta et al., 2011). Rather than base-pair mutations, NHEJ primarily leads to either single base-pair insertions, or multiple base-pair deletions (Gong et al., 2005; Aniukwu et al., 2008). We did not explicitly test for an association between FQ-R mutations and the frequency of INDELs due to the current limitations of bioinformatic programs to confidently identify INDELs along the *Mtb* chromosome using whole genome sequencing data. Further work is therefore required to elucidate whether FQ-R mutations lead to an increased frequency of INDELs in *Mtb*. This may be of critical importance, as multiple anti-TB drugs can have INDELs-based DR mechanisms, including the current first-line drugs INH and pyrazinamide (Gygli et al., 2017).

Overall, we showed that epistasis can modulate the in vitro evolution of DR in *Mtb*. Furthermore, we observed that FQ-R mutations had limited impact on the genetic diversity present in *Mtb*. Therefore, FQ exposure may not lead to more adaptive *Mtb* populations. Understanding how *Mtb* populations evolve in the presence of FQ pressure is of vital importance, as FQs are required to treat MDR-TB patients, and form an integral part in experimental first-line treatment regimen against drug-susceptible *Mtb*. Thus, our work highlights the potential usefulness of evolutionary studies in predicting the effect of new treatment regimens on natural populations of bacterial pathogens.

6 General Discussion

6.1 Synopsis of Main Findings

Fluoroquinolones (FQs) are important components in current treatment regimens against multidrug-resistant TB (MDR-TB) (Van Deun et al., 2010; Takiff et al., 2011; Nunn et al., 2019; WHO, 2019b), and in experimental treatment regimens against drug-susceptible TB (Gillespie et al., 2014; Merle et al., 2014; Jindani et al., 2014; Vjecha et al., 2018; Imperial et al., 2018). Investigations into the evolution of FQ-R in *M. tuberculosis* (*Mtb*) populations may provide insights in maintaining the current and potential efficacy of FQs.

Bacterial genetics have been shown to modulate the evolution of resistance to multiple anti-TB drugs (Gagneux et al., 2006a; Gagneux et al., 2006c; Zaczek et al., 2009; Fenner et al., 2012; Ford et al., 2013). Meanwhile, antimicrobial resistance (AMR) phenotypes, such as transmission potential and resistance levels, may differ depending on the given AMR mutation present (Gagneux et al., 2006c; Angst et al., 2013; Borrell et al., 2013; Vogwill et al., 2014; Vogwill et al., 2016; Huseby et al., 2017; Hughes et al., 2017). The main Aims of this study were to investigate whether bacterial genetics modulated FQ-resistance (FQ-R) in *Mtb*, and determine whether FQ-R mutations themselves could modulate the evolutionary trajectory of *Mtb* populations. In our first objective, we used extensive in vitro assays coupled with analysis of publicly available *Mtb* genomic sequences isolated from clinical strains to demonstrate that the genetic variation in *Mtb* can substantially modulate the frequency of FQ-R emergence and the mutational profiles for FQ-R. Furthermore, we observed that the FQ-R phenotypes were dependent on both the given FQ-R mutation present and the *Mtb* genetic background in which the mutation was present. For our second objective, we adapted the model of drug-resistance (DR) evolution developed by Ford et al. (2013) to test the relative contributions of three bacterial factors in determining the strain-dependent FQ-R frequencies in *Mtb*: DNA mutation rates, mutational target sizes for FQ-R, and fitness effects of FQ-R mutations. Our simulations provided two, non-exclusive conclusions. First, bacterial factors not studied here may play a large role in determining FQ-R frequencies. Second, a new model of DR evolution is required to explain the strain-dependent variation in FQ-R frequencies in *Mtb*. For our last objective, we again used in vitro assays coupled with population genomics analysis to test whether FQ-R mutations could confer a hypermutator phenotype in *Mtb* and consequently increase the levels of genetic diversity present in *Mtb* populations. Our in vitro results show that FQ-R may associate with increased frequencies of further DR acquisition; this association is dependent on both the FQ-R mutation present, and in which *Mtb* genetic background the FQ-R mutation was present. However, population genomics analysis showed no association between FQ-R mutations and increased genetic diversity, suggesting that FQ-R mutations have limited impact on the genetic diversity present in *Mtb* populations.

We next discuss the general limitations of our study. Considering these limitations, we then discuss the contributions of our findings to the study of FQ-R evolution in *Mtb*, to the study of pathogen genetics on DR evolution in general, and to public health. Throughout this chapter, we discuss potential research avenues that would be relevant to the given section, as well as provide additional research avenues that were not previously discussed but can also build upon our work.

6.2 General Limitations

Our study is limited by the fact that pathogen phenotypes in vitro may differ from pathogen phenotypes within patients. For example, our in vitro assays generally use nutrient rich medium that allows for optimal bacterial growth (see Methods of Chapters 3 and 5). However, bacterial growth in patients is generally curtailed by immune system functions including nutritional immunity (Berney et al., 2017), innate immunity (Liu et al., 2017), and adaptive immunity (Jasenosky et al., 2015). Furthermore, the immune system itself has also been hypothesized to modulate the prevalence of DR observed (Read et al., 2001; Handel et al., 2009; Chiang et al., 2018). Thus, our measurements of FQ-R frequencies and phenotypes in in vitro populations of *Mtb* may not be fully representative of the FQ-R frequencies and phenotypes in *Mtb* populations residing in patients. Nevertheless, we observed that up to 51% of the variability in the clinical frequency of FQ-R *gyrA* mutation events can be attributed to how *Mtb* evolves in vitro (see Chapter 3; Castro et al., 2019). Furthermore, multiple studies (Gagneux et al., 2006c; Andersson et al., 2010; Maruri et al., 2012; Angst et al., 2013; Vogwill et al., 2014; Vogwill et al., 2016; Huseby et al., 2017), including published work from this Thesis (see Chapter 3; Castro et al., 2019), have shown strong positive associations between the fitness of a given DR mutation in vitro with its respective prevalence in the clinic. Thus, while in vitro conditions may not be perfect representatives of patient infections, the combination of extensive in vitro work coupled with population genomics analysis can aid in understanding how *Mtb* populations evolve FQ-R in patients and, in general, how DR evolves in pathogens.

The adapted mathematical model from our second objective shares the same inherent limitation as other theoretical models. Namely, designing models generally requires a balance between maximizing the amount of variability in the experimental results that can be explained by the model, while minimizing the amount of variables that are required in the model to explain the variability (Spicknall et al., 2013; Blanquart, 2019). If the experimental data that the model is built around is itself biased or controversial, then the model would share the same biases. We hypothesize that this limitation may play a crucial role in explaining why our adapted model could not recapitulate the strain-dependent in vitro FQ-R frequencies when using experimentally-confirmed or inferred values for DNA mutation rates, mutational target sizes for FQ-R, and the fitness effects of FQ-R mutations (see Chapter 4). The original model developed by Ford et al. that we adapted for this study was designed based on their results suggesting that L2 Beijing strains have a hypermutator phenotype (Ford et al., 2013). Therefore, the Ford et al. model may inherently set mutation rates to have a higher relative contribution in

determining the DR frequencies than other bacterial factors. This suggests that our adapted model is also biased towards providing a higher relative contribution of mutation rates in determining FQ-R frequencies in vitro, while minimizing the relative contribution of mutational target sizes and fitness effects. An alternative, non-exclusive hypothesis is that there may be other bacterial factors relevant in determining FQ-R frequencies which we have not measured in vitro.

Our population genomics analyses are limited by the fact that most of the *Mtb* genomes we analyzed were sampled by convenience due to their public availability. While we controlled for differences in sequence analysis by reanalyzing all genomes using our standard sequence analysis pipeline (see Methods of Chapters 3 and 5), we cannot control for differences in DNA isolation techniques or sequencing methods between different studies. Differences in DNA isolation, processing, and sequencing may lead to differences in the SNPs that can be identified from genomic data, which can skew the level of genetic diversity present in a given *Mtb* sample (Meehan et al., 2019; Colman et al., 2019). Furthermore, the sequencing depth provided by the publicly genomes lacks the resolution to confidently identify minor variants in the population (*i.e.* variable SNPs that represent less than 10% of the population). Identifying minor variants in a population requires deep sequencing of DNA samples, which imposes a substantial financial cost (Trauner et al., 2017; Colman et al., 2019; Schwarze et al., 2019). Identifying minor variants may be important for studying FQ-R evolution. For instance, we may be underestimating the amount of genetic diversity present during our study outlined in Chapter 5. Furthermore, heteroresistance, defined as the co-existence of both wild-type and DR strains in the same infecting pathogen population, can be common in FQ-R evolution (Zhang et al., 2012; Eilertson et al., 2014; Trauner et al., 2017; Rigouts et al., 2019). Identification of low frequency FQ-R *Mtb* variants in patients may aid in designing personalized treatment regimens, which may reduce the overall prevalence of FQ-R *Mtb* strains. Thus, more work is required to determine whether minor variants have an impact on studying FQ-R evolution in *Mtb*.

6.3 Aspects in the Molecular Evolution of Fluoroquinolone-Resistance in *M. tuberculosis*

Considering the limitations listed above, we nevertheless observed robust signals that provide insights into the molecular evolution of FQ-R in *Mtb*. Specifically, we showed significant differences in FQ-R phenotypes depending on the FQ-R mutations present. For instance, *gyrA* G88C mutation conferred a large fitness cost irrespective of the *Mtb* genetic background it occurred in (see Chapter 3; Castro et al., 2019). This was as expected, as the *gyrA* G88C mutation is one of the least prevalent FQ-R mutation observed in the clinic (Maruri et al., 2012; Castro et al., 2019). The cost of G88C may be due to two features resulting from substituting in a cysteine in lieu of a glycine. Firstly, cysteine residues are larger than glycine residues. Secondly, cysteine residues can also interact with other cysteine residues to create disulfide bridges (Yen et al., 2000). The resulting change in residue size and/or interaction potential conferred by a *gyrA* G88C mutation may severely destabilize the structure and

function of DNA gyrase (Piton et al., 2010; Aldred et al., 2016; Blower et al., 2016; Pandey et al., 2018). Regardless of how the cost is inferred by the *gyrA* G88C mutation, designing new FQs types that can predominantly select for the *gyrA* G88C mutation may reduce the frequency of FQ-R observed in *Mtb*.

We also observed that *gyrA* D94G and A90V mutations conferred little or no fitness costs in vitro, and were the most prevalent FQ-R mutations both in vitro and in the clinic (see Chapter 3; Castro et al., 2019). Previous studies have shown that the D94G and A90V mutations confer different structural changes to DNA gyrase (Blower et al., 2016; Pandey et al., 2018). Therefore, we hypothesize that *gyrA* D94G and *gyrA* A90V confer low to no fitness costs through different biochemical and/or biological mechanisms. This potential difference in the causative mechanism of their respective fitness effects may be modulated by the *Mtb* genetic background. Indeed, we observed that *gyrA* D94G and *gyrA* A90V have different fitness effects in vitro depending on the *Mtb* strain they were present in, and had lineage-specific differences in clinical prevalence.

Interestingly, while *gyrA* D94A was the third-most prevalent in the clinic, *gyrA* D94N was the third-most prevalent mutation in vitro (see Chapter 3; Castro et al., 2019). Although we have no fitness measurements for the *gyrA* D94A mutation to compare with the fitness of *gyrA* D94N, out of 680 FQ-R colonies isolated during the in vitro mutational profile assay, only 3 *gyrA* D94A mutations were observed compared to 94 *gyrA* D94N mutations. This suggests that while *gyrA* D94A may confer a higher cost than *gyrA* D94N in vitro, the opposite is true in patients. Differences in the in vitro and clinical prevalence of DR mutations have been shown for other anti-TB drugs. For example, while the *katG* S315T is the most prevalent isoniazid-resistance (INH-R) mutation in the clinic (Gagneux et al., 2006a; Fenner et al., 2012; Casali et al., 2014; Manson et al., 2017; Wollenberg et al., 2017), it is rarely observed in vitro (Bergval et al., 2009; Brossier et al., 2016). One hypothesis to explain this is that *katG* S315T has a fitness advantage in the presence of host immune system pressure (Brossier et al., 2016). The same phenomenon may govern the higher than expected frequency of *gyrA* D94A in the clinic when compared to its in vitro prevalence. To test this hypothesis, studies that explore whether *gyrA* D94A mutations have higher fitness in infection models versus in vitro are required.

An alternate hypothesis for *gyrA* D94A being more prevalent than expected is that *gyrA* D94A interacts epistatically with other DR mutations. We hypothesize this due to two reasons. Firstly, all the genomes we surveyed for FQ-R mutations had both INH-R and rifampicin-resistance (RIF-R) mutations. Secondly, FQ-R mutations have been shown to confer different fitness effects depending on the presence of other DR mutations (Borrell et al., 2013; Huseby et al., 2017; Sun et al., 2018). Therefore, *gyrA* D94A mutation may confer lower fitness costs than *gyrA* D94N when in the presence of INH-R and/or RIF-R mutations.

6.4 Aspects in the Study of Pathogen Genetics modulating Drug-Resistance Evolution

We observed that the *Mtb* genetic background modulates multiple aspects in FQ-R evolution (see Chapters 3 and 5). This is in line with multiple studies showing *Mtb* lineages associating with differences DR evolution (Borrell et al., 2009; Fenner et al., 2012; Casali et al., 2014; Coscolla et al., 2014; Merker et al., 2015; Eldholm et al., 2016; Gagneux, 2018). However, we found that while FQ-R mutational profiles in the clinic associated with specific *Mtb* lineages, both in vitro FQ-R mutational profiles and frequencies associated with specific *Mtb* strains and not per lineage (see Chapter 3). This may be due to the use of only three strains per lineage for the in vitro studies, which may not be enough to delineate lineage-specific distributions for FQ-R phenotypes. Furthermore, the three representative *Mtb* strains per lineage were chosen to represent the greatest genetic diversity within a given lineage (Borrell et al., 2019); this decision may increase the likelihood of phenotyping *Mtb* strains with outlier phenotypes within a given lineage. Thus, the use of more strains per lineage may provide greater resolutions to delineate lineage-specific distributions for a given FQ-R phenotype.

This Thesis also highlights the role of the genetic background in modulating DR evolution in general. Indeed, the genetic background has been shown to modulate DR evolution not only in bacterial pathogens (Angst et al., 2013; Vogwill et al., 2014; Vogwill et al., 2016; Huseby et al., 2017; Apjok et al., 2019), but in eukaryotic organisms such as *Leishmania donovani* (Decuypere et al., 2012) and in yeast as well (Mullis et al., 2018). Thus, the epistasis between DR mutations and the genetic variation present in any given organism appears to be a major factor in the DR evolution.

6.5 Public Health Relevance

The role of *Mtb* genetics influencing FQ-R evolution may have a clinical impact. Specifically, considering the non-random geographic distribution of *Mtb* genotypes (Comas et al., 2010; Gagneux, 2018), different *Mtb* strains associating with different FQ-R frequencies in vitro may translate into regional differences in FQ-R prevalence. For instance, the Lineage 1 "N0157" strain is associated with high FQ-R frequencies in vitro (see Chapter 3; Castro et al., 2019), and belongs to the "Manila" sublineage, the predominant *Mtb* genotype in the Philippines and in TB patients of Filipino descent (Douglas et al., 2003; Gagneux et al., 2006b; Wan et al., 2017). We hypothesize that FQ-R prevalence may be higher in the Philippines and in Filipinos compared to other countries or ethnic groups when using FQs as a first-line anti-TB drug. However, the frequency of FQ-R in more Manila sub-lineage strains should be measured before this hypothesis can be supported. Nevertheless, considering that the Philippines has the second-highest incidence of TB globally (WHO, 2019a), future work in FQ-R evolution should address this question.

Differences in the molecular evolution of FQ-R may also lead to differential treatment outcomes when using FQs as a first-line anti-TB drug. For instance, while both *gyrA* D94G and A90V muta-

tions are the most frequent FQ-R mutations in vitro and in the clinic (see Chapter 3; Castro et al., 2019), *gyrA* D94G associated with poorer treatment outcomes compared to *gyrA* A90V in MDR-TB patients (Rigouts et al., 2016; Farhat et al., 2017). Our results showed that while *gyrA* A90V was most prevalent in high-frequency FQ-R *Mtb* strains like the Lineage 1 Manila N0157 strain, *gyrA* D94G was most prevalent in the other strains (see Chapter 3). Furthermore, the *gyrA* D94G and A90V mutations associated with differences in further DR acquisition, with *gyrA* D94G associating with increased frequencies of further DR while *gyrA* A90V did not (see Chapter 5). This may lead to regional differences in treatment outcomes. For instance, regions where *Mtb* strains more likely to acquire *gyrA* A90V may associate with better treatment outcomes and less further DR mutations acquired compared to regions where *Mtb* strains would more likely acquire *gyrA* D94G. Future studies could therefore address the potential impact of FQ-R molecular evolution differences on treatment outcomes.

6.6 Potential Future Directions

In addition to the open questions posited in the preceding subsections and chapter discussions, we highlight two other potential avenues of research for FQ-R evolution in *Mtb*. Firstly, the fitness of any given genotype can be modulated by the environmental conditions it is present in (Wiesch et al., 2010; Huseby et al., 2017; Mullis et al., 2018). While we measured the fitness of FQ-R mutants in the absence of drug pressure, the fitness effects of FQ-R mutations may differ if FQs or other drugs are present. Measuring the fitness effects of FQ-R mutations in the presence of drugs may provide insights into the potential evolutionary trajectories of FQ-R mutants (Huseby et al., 2017). The fitness of FQ-R mutants in the presence of bedaquiline, pretomanid, or pyrazinamide are of particular interest, as a current Phase III clinical trial is testing the use of these three drugs in combination with FQs as a potential first-line anti-TB regimen (Vjecha et al., 2018).

Secondly, the further acquisition of secondary mutations may alleviate any initial fitness costs of FQ-R mutations. Such "compensatory" mutations have been observed alleviating costs of RIF-R in *Mtb* and *Salmonella* (Comas et al., 2012; Brandis et al., 2013; Song et al., 2014), and have been associated with ongoing transmission of RIF-R *Mtb* strains in the clinic (Vos et al., 2013; Merker et al., 2018). Compensatory mutations for FQ-R have yet to be identified; however, their existence may lead to more transmissible FQ-R *Mtb* strains.

6.7 Conclusion

In general, this Thesis highlights the importance of bacterial genetics on AMR evolution. Specifically, we provide the TB community, and the AMR community more broadly, with the first evidence showing that the *Mtb* genetic background modulates FQ-R frequencies and phenotypes. Further, we provide the TB community with the first in vitro fitness measurements for FQ-R mutations, one of the most extensive phenotyping of FQ-R mutants, and one of the largest in vitro and clinical muta-

tional profiling for FQ-R in *Mtb*. More generally, our study illustrates how the genetic variation in bacterial populations affects the emergence, the maintenance, and the phenotypes of clinically relevant chromosomal AMR mutations. The work presented in this Thesis, coupled with results from published studies, provides compelling evidence that the epistasis between pathogen genetics and AMR mutations is a major factor in AMR evolution. Our work may provide insights into how to maximize the time-span in which antimicrobials remain effective in the clinic.

7 Supplementary Information

7.1 Supplementary Information for Chapters 3 and 4

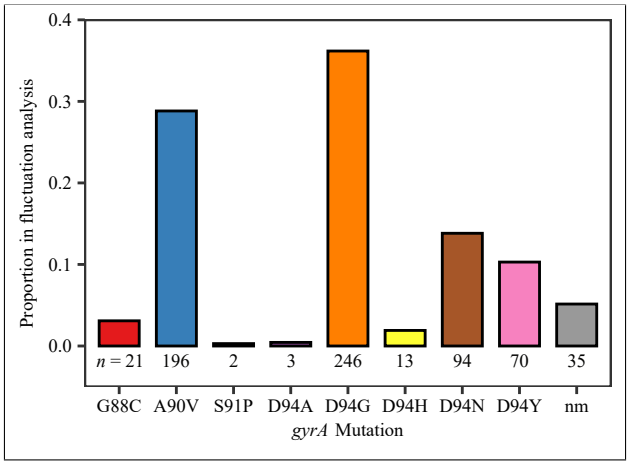


Figure 7.1: Proportion of each *gyrA* mutation after sequencing of the QRDR of *gyrA* in 680 ofloxacin-resistant colonies from the fluctuation analysis performed in Figure 3.1A (nm = no identified QRDR *gyrA* mutations).
The numbers of colonies with the given *gyrA* mutation are reported directly below each respective column.

Table 7.1: Classification of *M. tuberculosis* strains used for in vitro assays

Strain	Lineage	Sub-lineage	Alternate Sub-lineage/Strain Nomenclatures
N0069	L1	L1.1.1	EAS042 (Hershberg et al., 2008)
N0072	L1	L1.1.2	EAS053 (Hershberg et al., 2008)
N0157	L1	L1.2.1	Manila; T92 (Tsolaki et al., 2004)
N0052	L2	L2.2.2	Beijing, Asia Central 1; 98_1833 (Hershberg et al., 2008; Coll et al., 2014; Shitikov et al., 2017)
N0145	L2	L2.2.1.1	Beijing, Pacific RD150; T67 (Tsolaki et al., 2004; Coll et al., 2014; Shitikov et al., 2017)
N0155	L2	L2.2.1	Beijing; T85 (Tsolaki et al., 2004; Gagneux et al., 2006c)
N0136	L4	L4.3.3	Beijing; Latin America-Mediterranean; T4 (Tsolaki et al., 2004; Coll et al., 2014; Stucki et al., 2016)
N1216	L4	L4.6.2.2	Cameroon (Coll et al., 2014; Stucki et al., 2016
N1283	L4	L4.2.1	Ural (Coll et al., 2014; Stucki et al., 2016

Lineage 1 = L1; Lineage 2 = L2; Lineage 4 = L4. *Mtb* lineage designations defined as in Comas et al., 2010; Gagneux, 2018.

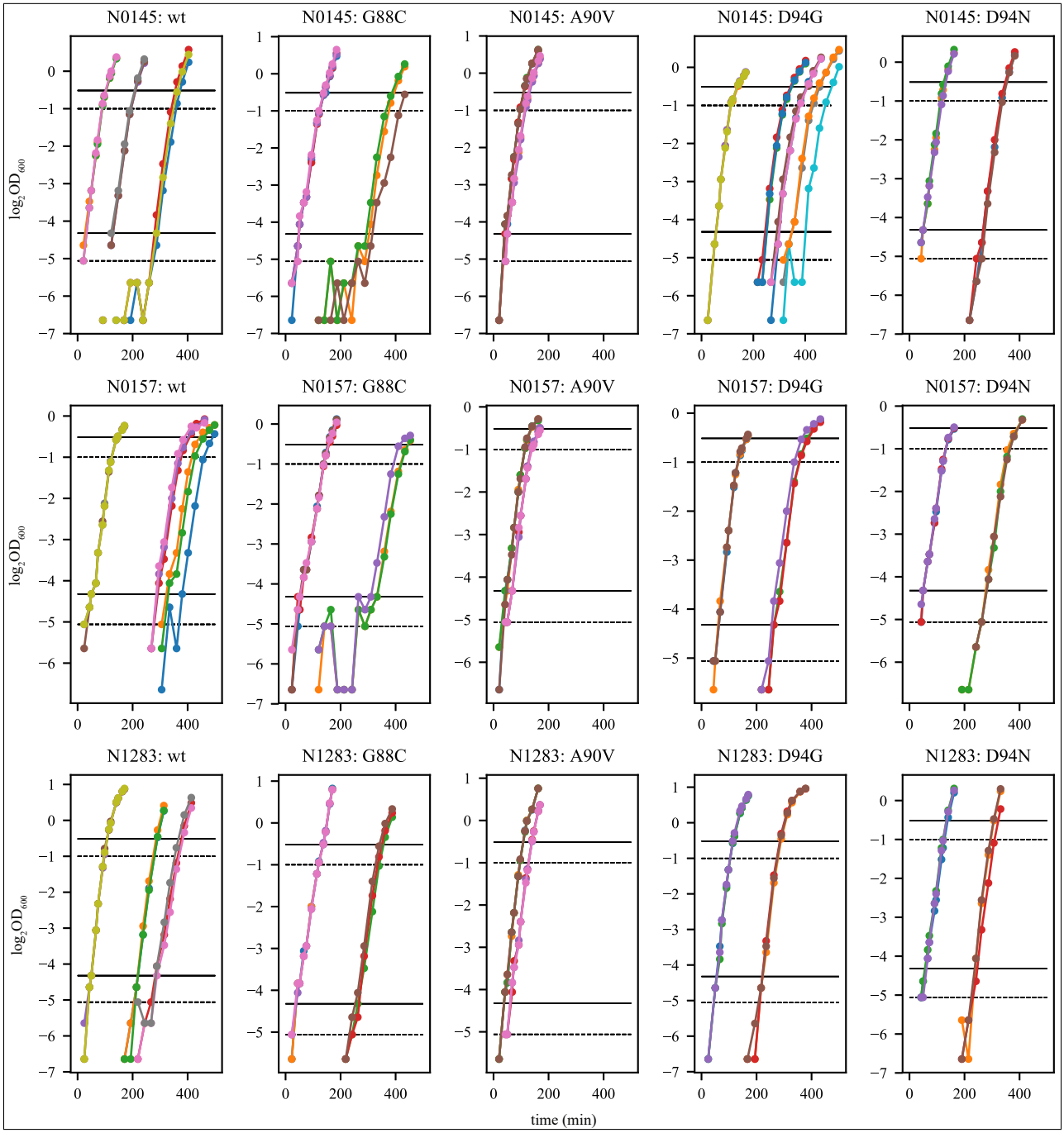


Figure 7.2: Growth profiles of *M. tuberculosis* strains in cell growth assays under antibiotic free conditions, with all OD_{600} values plotted (log_2-transformed).
Genetic background of *M. tuberculosis* strain and its corresponding *gyrA* mutation are presented above each respective plot. Coloured dots represent the measured log_2OD_{600} values at a given time (in minutes), with coloured lines connecting respective coloured dots. Different colours represent different replicates for each strain. For reference, black solid horizontal lines were plotted to denote non-transformed OD_{600} values of either 0.05 (lower line) or 0.70 (upper line), while black dashed lines denote OD_{600} values of either 0.03 (lower line) or 0.50 (upper line).

Table 7.2: Phylogenetic single nucleotide polymorphisms leading to missense DNA gyrase or DnaE mutations that are present in the genomic data of the nine drug-susceptible *M. tuberculosis* strains outlined in Supplementary Table 7.1

Gene	Gene Name	Amino Acid Substitution	Strain	Lineage
Rv0005	*gyrB*	M291I	N0069	L1
Rv0005	*gyrB*	M291I	N0072	L1
Rv0005	*gyrB*	M291I	N0157	L1
Rv0006	*gyrA*	A384V	N0069	L1
Rv0006	*gyrA*	A384V	N0072	L1
Rv0006	*gyrA*	A384V	N0157	L1
Rv0006	*gyrA*	G247S	N0136	L1
Rv0006	*gyrA*	K224E	N0072	L1
Rv0006	*gyrA*	P154R	N1216	L4
Rv1547	*dnaE1*	D316N*	N0069	L1
Rv1547	*dnaE1*	S898L*	N0052	L1
Rv3370c	*dnaE2*	C313W*	N0157	L1
Rv3370c	*dnaE2*	P814S*	N0052	L2

*These mutations do not occur in the polymerase and histidinol phosphatase-domain of DNA polymerase *dnaE*, and therefore are not confirmed to confer a hypermutator phenotype in *M. tuberculosis* (Rock et al., 2015; Baños-Mateos et al., 2017).

Table 7.3: Mutations present in the *rpsL* gene for 194 streptomycin-resistant colonies following fluctuation analysis on 100 μg/mL of streptomycin

Mutation	N0157	N1283	N0145
K43M	0	4	0
K43N	1	4	0
K43R	5	72	46
K43R, K10Q	0	1	0
K43T	9	13	13
K88E	0	3	0
K88R	0	5	0
nm	0	9	9

nm = streptomycin-resistant colonies with no mutations in *rpsL*.

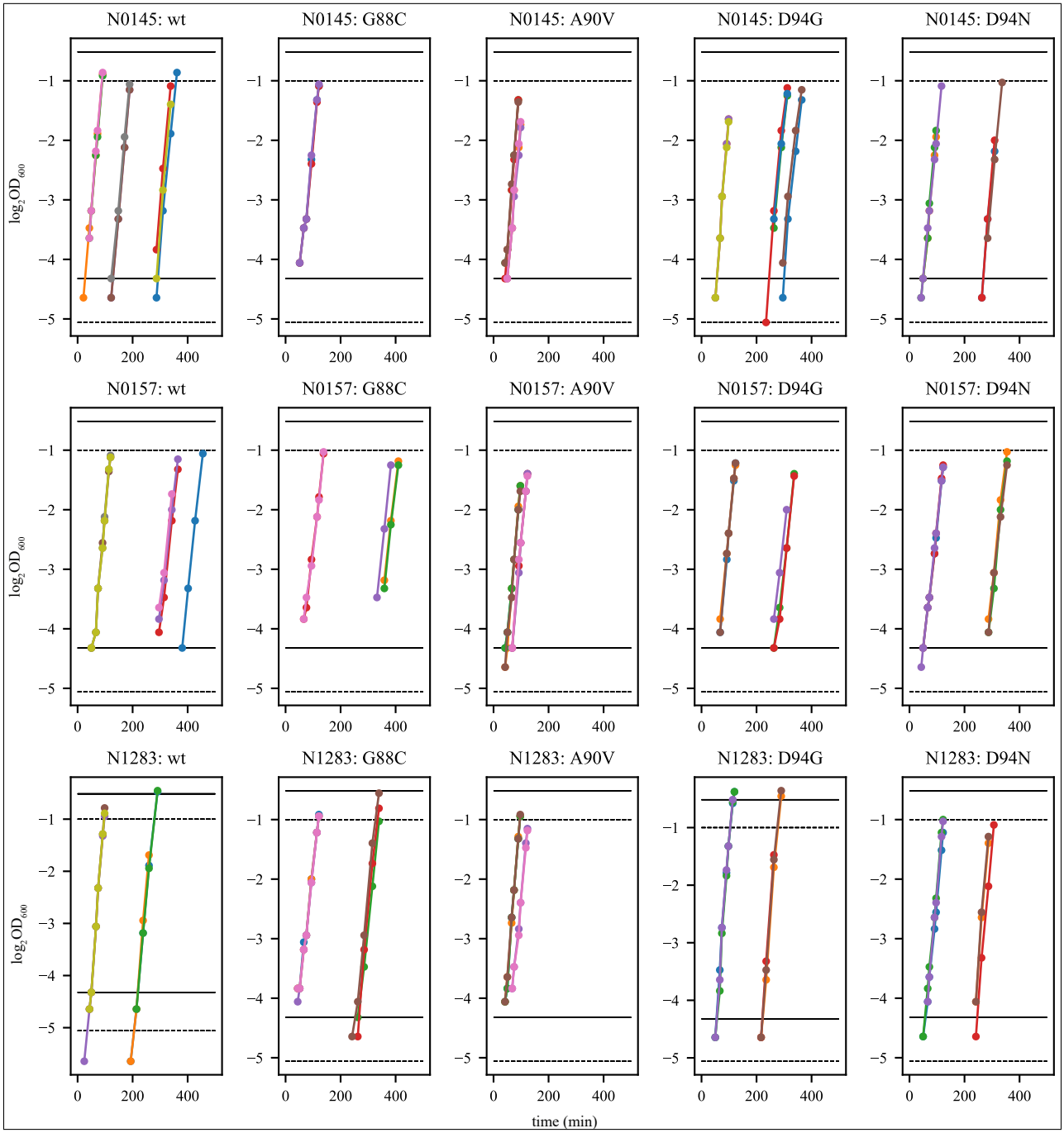


Figure 7.3: Growth profiles of *M. tuberculosis* strains in cell growth assays under antibiotic free conditions, with only measured OD_{600} values (log_2-transformed) present after filtering for exponential phase of growth.
Genetic background of *M. tuberculosis* strain and its corresponding *gyrA* mutation are presented above each respective plot. Exponential growth phase was defined as a set consecutive time-points where a linear relationship between log_2OD_{600} (defined by a Pearson's R^2 value $\geq$0.98) and time was present. Coloured dots represent the measured log2OD600 values at a given time (in minutes), with coloured lines connecting respective coloured dots. Different colours represent different replicates for each strain. For reference, black solid horizontal lines were plotted to denote non-transformed OD_{600} values of either 0.05 (lower line) or 0.70 (upper line), while black dashed lines denote OD_{600} values of either 0.03 (lower line) or 0.50 (upper line).

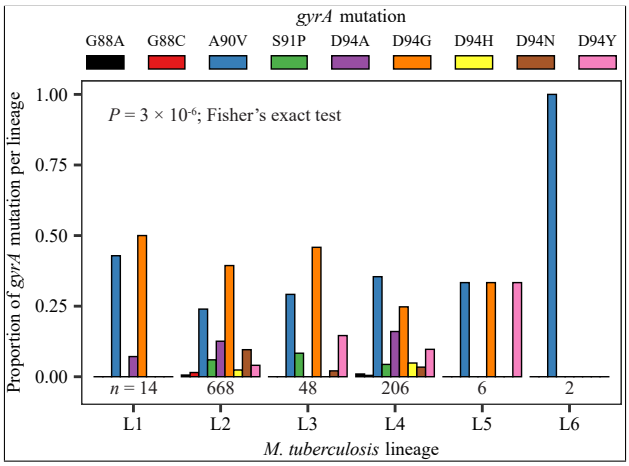


Figure 7.4: Mutational profile for all (fixed and variable) fluoroquinolone-resistance *gyrA* mutations is lineage-specific in clinical isolates of *M. tuberculosis*.
An initial dataset of 3,450 genomes with confirmed MDR-TB mutations were surveyed. 854 genomes were identified as fluoroquinolone-resistant, with 848 of these genomes containing gyrA mutations. If genomic data from a single *M. tuberculosis* clinical isolate contained multiple fluoroquinolone-resistance *gyrA* mutation, each mutation was counted once (n = 944). Number of genomes analyzed per lineage is presented underneath each respective bar graph. *Mtb* lineage designations defined as in Comas et al., 2010; Gagneux, 2018.

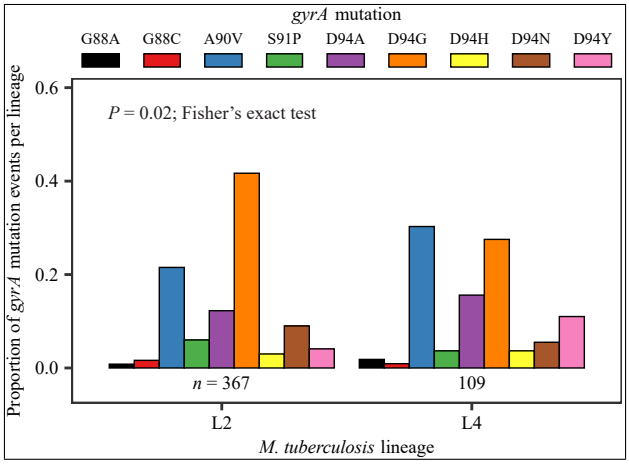


Figure 7.5: Variation in the frequency of mutation events per fluoroquinolone-resistance (FQ-R) *gyrA* mutation amongst clinical isolates of *M. tuberculosis* belonging to either L2 or L4 lineages.
Mutation events per FQ-R *gyrA* mutation were enumerated from an initial dataset of 3,450 genomes with confirmed MDR-TB mutations. Genomes were defined as belonging to a given transmission cluster by using a pairwise genetic distance threshold of 12 single nucleotide polymorphisms average as a cut-off. Each unique and fixed FQ-R *gyrA* mutation present per transmission cluster, as well as each fixed FQ-R *gyrA* mutation present in non-clustered genomes, were counted as independent mutation events. The number of FQ-R *gyrA* mutation events per lineage is presented underneath each respective bar graph. *Mtb* lineage designations defined as in Comas et al., 2010; Gagneux, 2018.

Table 7.4: Mutations in the QRDR of *gyrA* for 680 ofloxacin-resistant colonies following fluctuation analysis on 4 μg/mL of ofloxacin

Strain	G88C	A90V	S91P	D94A	D94G	D94H	D94N	D94Y	nm	Strain Total
N0157	1	43	0	0	9	1	7	1	2	64
N0072	2	112	0	2	37	0	1	11	7	172
N0052	13	4	0	0	39	2	18	15	21	112
N0155	2	0	0	0	51	0	14	14	3	84
N1283	1	12	1	0	36	7	14	19	0	90
N0136	0	24	1	1	31	3	13	9	0	82
N1216	1	0	0	0	27	0	16	1	0	45
N0069	1	1	0	0	13	0	4	0	2	21
N0145	0	0	0	0	3	0	7	0	0	10
Mutation Total	21	196	2	3	246	13	94	70	35	680

nm = ofloxacin-resistant colonies with no mutations in the QRDR region of *gyrA*. Strains are ordered top to bottom based on their frequency of OFX-resistance at 4 μg/mL OFX as shown in Figure 3.1A.

Table 7.5: Mutations in the QRDR of *gyrB* for 590 ofloxacin-resistant colonies following fluctuation analysis on 4 μg/mL of ofloxacin

Strain	E454K	D461H	nm	Strain Total
N0157	0	1	41	22
N0072	0	0	155	155
N0052	1	0	101	102
N0155	0	0	66	7
N1283	0	0	69	46
N0136	0	0	81	69
N1216	0	0	46	81
N0069	0	0	22	42
N0145	0	0	7	66
Mutation Total	1	1	588	590

nm = ofloxacin-resistant colonies with no mutations in the QRDR region of *gyrB*. Strains are ordered top to bottom based on their frequency of OFX-resistance at 4 μg/mL OFX as shown in Figure 3.1A.

Table 7.6: Ofloxacin MIC estimates for *gyrA* mutant strains and their respective parental strain

Strain	gyrA Mutation	Genetic Background (Parental Strain)	Ofloxacin MIC (μg/mL)	Normalized Ofloxacin MIC*
N0157	wt	—	2.00	1.00
N3661	G88C	N0157	31.60	15.80
N2034	A90V	N0157	10.00	5.00
N2036	D94G	N0157	20.00	10.00
N2035	D94N	N0157	20.00	10.00
N1283	wt	—	0.60	1.00
N2508	G88C	N1283	12.60	21.00
N2505	A90V	N1283	4.00	6.67
N3915	D94G	N1283	12.60	21.00
N2507	D94N	N1283	12.60	21.00
N0145	wt	—	0.50	1.00
N3659	G88C	N0145	39.80	79.60
N2847	A90V	N0145	3.20	6.40
N1893	D94G	N0145	10.00	20.00
N1895	D94N	N0145	10.00	20.00

MIC estimates for Ofloxacin based on fitting of a Hill curve to the distribution of fluorescence in an Alamar Blue assay (Franzblau et al., 1998). MIC is defined as the ofloxacin concentration where fitted Hill curve showed a $\geq$95% reduction in fluorescence. *Normalized Ofloxacin MIC is calculated by taking the ofloxacin MIC of a given *M. tuberculosis* strain and dividing it by the ofloxacin MIC of its respective wild-type parental strain; Normalized Ofloxacin MICs for each wild-type parental strain are therefore equal to 1.00.

Table 7.7: In vitro fitness of *M. tuberculosis* strains based on cell growth assays in antibiotic-free conditions

Strain	*gyrA* Mutation	Genetic Background (Parental Strain)	Growth Rate (GR)	GR: Lower 95%	GR: Upper 95%	Generation Time (in hours)	Relative Fitness (RF)	RF: Lower 95%	RF: Upper 95%	*P*
N0157	wt	—	0.045	0.044	0.047	22.22				
N3661	G88C	N0157	0.038	0.034	0.042	26.36	0.844	0.773	0.894	<0.001*
N2034	A90V	N0157	0.052	0.048	0.057	19.23	1.156	1.091	1.213	<0.001*
N2036	D94G	N0157	0.044	0.04	0.048	22.73	0.978	0.909	1.021	0.354
N2035	D94N	N0157	0.042	0.038	0.046	23.81	0.933	0.864	0.979	0.009*
N1283	wt	—	0.061	0.059	0.064	16.40				
N2508	G88C	N1283	0.042	0.037	0.047	23.81	0.689	0.627	0.734	<0.001*
N2505	A90V	N1283	0.054	0.048	0.06	18.52	0.885	0.814	0.938	<0.001*
N3915	D94G	N1283	0.062	0.056	0.068	16.13	1.016	0.949	1.062	0.638
N2507	D94N	N1283	0.052	0.046	0.058	19.23	0.852	0.780	0.906	<0.001*
N0145	wt	—	0.053	0.05	0.055	18.87				
N3659	G88C	N0145	0.044	0.038	0.05	22.73	0.830	0.760	0.909	<0.001*
N2847	A90V	N0145	0.058	0.051	0.063	17.24	1.094	1.020	1.145	0.016*
N1893	D94G	N0145	0.05	0.044	0.056	20.00	0.943	0.880	1.018	0.107
N1895	D94N	N0145	0.051	0.044	0.056	19.61	0.962	0.880	1.018	0.206

The growth rate of a particular *M. tuberculosis* strain was defined as the slope during exponential phase of bacterial growth, with the exponential phase of bacterial growth defined as where a $\log_2$-linear relationship existed between OD_{600} and time using a Pearson's R^2 value ≥ 0.98 as the threshold. Generation times were calculated by taking the inverse of the calculated growth rate. The relative fitness of a given *gyrA* mutant was defined by taking its growth rate and dividing it by the growth rate of its respective wild-type ancestor.

Table 7.8: Number of publicly available genomes from *M. tuberculosis* clinical isolates used to survey the mutational profile for fluoroquinolone-resistance

Mtb lineage	No. of MDR-TB Genomes	No. of MDR-TB + FQ-R Genomes
Lineage 1	109	13
Lineage 2	1,903	597
Lineage 3	151	44
Lineage 4	1,261	196
Lineage 5	18	2
Lineage 6	8	2
Total	3,450	854

Mtb = *M. tuberculosis*; MDR-TB = multidrug-resistant tuberculosis, defined as *Mtb* genomes that have both an isoniazid and a rifampicin-resistance mutation; FQ-R = fluoroquinolone-resistant, defined as *Mtb* genomes that have a fluoroquinolone-resistance mutation. Multiple drug-resistance mutations present in the genomic data from a single *Mtb* clinical isolate is possible (classified as "variable," and therefore not "fixed" for drug-resistance mutations); if a genome contained multiple drug-resistance mutations, then the genome is simply counted as MDR-TB or FQ-R once. *Mtb* lineage designations defined as in Comas et al., 2010; Gagneux, 2018.

Table 7.9: Frequency of all (fixed and variable) fluoroquinolone-resistance mutations from sample set of 3,450 MDR-TB genomes.

Mutation	L1	L2	L3	L4	L5	L6	Mutation Total
gyrA G88A	0	4	0	2	0	0	6
gyrA G88C	0	10	0	1	0	0	11
gyrA A90V	6	160	14	73	2	2	257
gyrA S91P	0	40	4	9	0	0	53
gyrA D94A	1	84	0	33	0	0	118
gyrA D94G	7	263	22	51	2	0	345
gyrA D94H	0	16	0	10	0	0	26
gyrA D94N	0	64	1	7	0	0	72
gyrA D94Y	0	27	7	20	2	0	56
gyrB D461N	0	3	0	2	0	0	5
gyrB N499D	0	1	0	0	0	0	1
Lineage Total	14	672	48	208	6	2	950

If a genome was classified as "variable" for fluoroquinolone-resistance mutations, each *gyrA* or *gyrB* mutation present was counted once. Lineage 1 = L1, Lineage 2 = L2, Lineage 3 = L3, Lineage 4 = L4, Lineage 6 = L5, Lineage 6 = L6. Mtb lineage designations defined as in Comas et al., 2010; Gagneux, 2018.

Table 7.10: Frequency of fixed fluoroquinolone-resistance mutations from sample set of 3,450 MDR-TB genomes.

Mutation	L1	L2	L3	L4	L5	L6	Mutation Total
gyrA G88A	0	3	0	2	0	0	5
gyrA G88C	0	8	0	1	0	0	9
gyrA A90V	4	114	9	61	0	2	190
gyrA S91P	0	26	4	7	0	0	37
gyrA D94A	0	70	0	26	0	0	96
gyrA D94G	6	200	16	39	0	0	261
gyrA D94H	0	12	0	4	0	0	16
gyrA D94N	0	48	1	6	0	0	55
gyrA D94Y	0	19	5	17	0	0	41
gyrB D461N	0	3	0	2	0	0	5
gyrB N499D	0	1	0	0	0	0	1
Lineage Total	10	504	35	165	0	2	716

If a genome was classified as "variable" for fluoroquinolone-resistance mutations, each *gyrA* or *gyrB* mutation present was counted once. Lineage 1 = L1, Lineage 2 = L2, Lineage 3 = L3, Lineage 4 = L4, Lineage 6 = L5, Lineage 6 = L6. Mtb lineage designations defined as in Comas et al., 2010; Gagneux, 2018.

Table 7.11: Frequency of mutation events per fluoroquinolone-resistance gyrA mutation from an initial sample set of 3,450 MDR-TB genomes.

Mutation	L1	L2	L3	L4	L5	L6	Mutation Events Total
gyrA G88A	0	3	0	2	0	0	5
gyrA G88C	0	6	0	1	0	0	7
gyrA A90V	4	79	9	33	0	2	127
gyrA S91P	0	22	3	4	0	0	29
gyrA D94A	0	45	0	17	0	0	62
gyrA D94G	3	153	15	30	0	0	201
gyrA D94H	0	11	0	4	0	0	15
gyrA D94N	0	33	1	6	0	0	40
gyrA D94Y	0	15	3	12	0	0	30
Lineage Total	7	367	31	109	0	2	516

Only "fixed" fluoroquinolone-resistance *gyrA* mutation events were enumerated here. Notably, no "fixed" mutations were observed in Lineage 5 (L5) strains. Lineage 1 = L1, Lineage 2 = L2, Lineage 3 = L3, Lineage 4 = L4, Lineage 6 = L6. Mtb lineage designations defined as in (Comas et al., 2010; Gagneux, 2018).

Bibliography

Ahmad, N et al. (2018). "Treatment correlates of successful outcomes in pulmonary multidrug-resistant tuberculosis: an individual patient data meta-analysis." *The Lancet* 392.10150, pp. 821–834.

Aldred, KJ, TR Blower, RJ Kerns, JM Berger, and N Osheroff (2016). "Fluoroquinolone interactions with *Mycobacterium tuberculosis* gyrase: Enhancing drug activity against wild-type and resistant gyrase." *Proceedings of the National Academy of Sciences* 113.7, E839–E846.

Aldred, KJ, RJ Kerns, and N Osheroff (2014). "Mechanism of Quinolone Action and Resistance." *Biochemistry* 53.10, pp. 1565–1574.

Alene, KA, H Yi, K Viney, ES McBryde, K Yang, L Bai, DJ Gray, ACA Clements, and Z Xu (2017). "Treatment outcomes of patients with multidrug-resistant and extensively drug resistant tuberculosis in Hunan Province, China." *BMC Infectious Diseases* 17.1, p. 573.

Alexander, KA, CE Sanderson, MH Larsen, S Robbe-Austerman, MC Williams, and MV Palmer (2016). "Emerging Tuberculosis Pathogen Hijacks Social Communication Behavior in the Group-Living Banded Mongoose (*Mungos mungo*)." *mBio* 7.3.

Alipanah, N, L Jarlsberg, C Miller, NN Linh, D Falzon, E Jaramillo, and P Nahid (2018). "Adherence interventions and outcomes of tuberculosis treatment: A systematic review and meta-analysis of trials and observational studies." *PLOS Medicine* 15.7, e1002595.

Alvarez-Uria, G, S Gandra, and R Laxminarayan (2016). "Poverty and prevalence of antimicrobial resistance in invasive isolates." *International Journal of Infectious Diseases* 52, pp. 59–61.

Andersson, DI and D Hughes (2010). "Antibiotic resistance and its cost: is it possible to reverse resistance?" *Nature Reviews Microbiology* 8.4, pp. 260–271.

Ängeby, KA, P Jureen, CG Giske, E Chryssanthou, E Sturegård, M Nordvall, AG Johansson, J Werngren, G Kahlmeter, SE Hoffner, and T Schön (2010). "Wild-type MIC distributions of four fluoroquinolones active against *Mycobacterium tuberculosis* in relation to current critical concentrations and available pharmacokinetic and pharmacodynamic data." *Journal of Antimicrobial Chemotherapy* 65.5, pp. 946–952.

Angst, DC and AR Hall (2013). "The cost of antibiotic resistance depends on evolutionary history in *Escherichia coli*." *BMC Evolutionary Biology* 13.1, p. 163.

Aniukwu, J, MS Glickman, and S Shuman (2008). "The pathways and outcomes of mycobacterial NHEJ depend on the structure of the broken DNA ends." *Genes & Development* 22.4, pp. 512–527.

Apjok, G, G Boross, Á Nyerges, G Fekete, V Lázár, B Papp, C Pál, and B Csörgő (2019). "Limited Evolutionary Conservation of the Phenotypic Effects of Antibiotic Resistance Mutations." *Molecular Biology and Evolution* 36.8, pp. 1601–1611.

Ayora, S, B Carrasco, PP Cárdenas, CE César, C Cañas, T Yadav, C Marchisone, and JC Alonso (2011). "Double-strand break repair in bacteria: a view from *Bacillus subtilis*." *FEMS Microbiology Reviews* 35.6, pp. 1055–1081.

Baharoglu, Z and D Mazel (2014). "SOS, the formidable strategy of bacteria against aggressions." *FEMS Microbiology Reviews* 38.6, pp. 1126–1145.

Balaban, NQ, J Merrin, R Chait, L Kowalik, and S Leibler (2004). "Bacterial Persistence as a Phenotypic Switch." *Science* 305.5690, pp. 1622–1625.

Baños-Mateos, S, AMMv Roon, UF Lang, SL Maslen, JM Skehel, and MH Lamers (2017). "High-fidelity DNA replication in *Mycobacterium tuberculosis* relies on a trinuclear zinc center." *Nature Communications* 8.1, pp. 1–10.

Bao, QS, YH Du, and CY Lu (2007). "Treatment outcome of new pulmonary tuberculosis in Guangzhou, China 1993–2002: a register-based cohort study." *BMC Public Health* 7.1, p. 344.

Barrick, JE and RE Lenski (2013). "Genome dynamics during experimental evolution." *Nature Reviews Genetics* 14.12, pp. 827–839.

Barter, DM, SO Agboola, MB Murray, and T Bärnighausen (2012). "Tuberculosis and poverty: the contribution of patient costs in sub-Saharan Africa – a systematic review." *BMC Public Health* 12.1, p. 980.

Baya, B, B Diarra, S Diabate, B Kone, D Goita, YdS Sarro, K Cohen, JL Holl, CJ Achenbach, M Tolofoudie, ACG Togo, M Sanogo, A Kone, O Kodio, D Dabitao, N Coulibaly, S Siddiqui, S Diop, W Bishai, S Dao, S Doumbia, RL Murphy, S Diallo, and M Maiga (2019). "Association of *Mycobacterium africanum* Infection with Slower Disease Progression Compared with *Mycobacterium tuberculosis* in Malian Patients with Tuberculosis." *The American Journal of Tropical Medicine and Hygiene*, tpmd190264.

Bergval, IL, ARJ Schuitema, PR Klatser, and RM Anthony (2009). "Resistant mutants of *Mycobacterium tuberculosis* selected *in vitro* do not reflect the *in vivo* mechanism of isoniazid resistance." *Journal of Antimicrobial Chemotherapy* 64.3, pp. 515–523.

Berney, M and L Berney-Meyer (2017). "*Mycobacterium tuberculosis* in the Face of Host-Imposed Nutrient Limitation." *Microbiology Spectrum* 5.3.

Blanquart, F (2019). "Evolutionary epidemiology models to predict the dynamics of antibiotic resistance." *Evolutionary Applications* 12.3, pp. 365–383.

Blower, TR, BH Williamson, RJ Kerns, and JM Berger (2016). "Crystal structure and stability of gyrase–fluoroquinolone cleaved complexes from *Mycobacterium tuberculosis*." *Proceedings of the National Academy of Sciences* 113.7, pp. 1706–1713.

Bolger, AM, M Lohse, and B Usadel (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." *Bioinformatics* 30.15, pp. 2114–2120.

Bonhoeffer, S, M Lipsitch, and BR Levin (1997). "Evaluating treatment protocols to prevent antibiotic resistance." *Proceedings of the National Academy of Sciences* 94.22, pp. 12106–12111.

Boritsch, EC, V Khanna, A Pawlik, N Honoré, VH Navas, L Ma, C Bouchier, T Seemann, P Supply, TP Stinear, and R Brosch (2016). "Key experimental evidence of chromosomal DNA transfer among selected tuberculosis-causing mycobacteria." *Proceedings of the National Academy of Sciences* 113.35, pp. 9876–9881.

Borrell, S and S Gagneux (2009). "Infectiousness, reproductive tness and evolution of drug-resistant *Mycobacterium tuberculosis*." *The International Journal of Tuberculosis and Lung Disease* 13.12, pp. 1456–1466.

Borrell, S, Y Teo, F Giardina, EM Streicher, M Klopper, J Feldmann, B Müller, TC Victor, and S Gagneux (2013). "Epistasis between antibiotic resistance mutations drives the evolution of extensively drug-resistant tuberculosis." *Evolution, Medicine, and Public Health* 2013.1, pp. 65–74.

Borrell, S, A Trauner, D Brites, L Rigouts, C Loiseau, M Coscolla, S Niemann, BD Jong, D Yeboah-Manu, M Kato-Maeda, J Feldmann, M Reinhard, C Beisel, and S Gagneux (2019). "Reference set of *Mycobacterium tuberculosis* clinical strains: A tool for research and product development." *PLOS ONE* 14.3, e0214088.

Bos, KI, KM Harkins, A Herbig, M Coscolla, N Weber, I Comas, SA Forrest, JM Bryant, SR Harris, VJ Schuenemann, TJ Campbell, K Majander, AK Wilbur, RA Guichon, DL Wolfe Steadman, DC Cook, S Niemann, MA Behr, M Zumarraga, R Bastida, D Huson, K Nieselt, D Young, J Parkhill, JE Buikstra, S Gagneux, AC Stone, and J Krause (2014). "Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis." *Nature* 514.7523, pp. 494–497.

Brandis, G and D Hughes (2013). "Genetic characterization of compensatory evolution in strains carrying *rpoB* Ser531Leu, the rifampicin resistance mutation most frequently found in clinical isolates." *Journal of Antimicrobial Chemotherapy* 68.11, pp. 2493–2497.

Brossier, F, E Cambau, E Tessier, V Jarlier, and W Sougakoff (2016). "The *in vitro* mechanisms of isoniazid and ethionamide resistance poorly reflect those *in vivo* in *Mycobacterium tuberculosis*." *Tuberculosis* 101, pp. 144–145.

Brynildsrud, OB, CS Pepperell, P Suffys, L Grandjean, J Monteserin, N Debech, J Bohlin, K Alfsnes, JOH Pettersson, I Kirkeleite, F Fandinho, MAd Silva, J Perdigao, I Portugal, M Viveiros, T Clark, M Caws, S Dunstan, PVK Thai, B Lopez, V Ritacco, A Kitchen, TS Brown, Dv Soolingen, MB O'Neill, KE Holt, EJ Feil, B Mathema, F Balloux, and V Eldholm (2018). "Global expansion of *Mycobacterium tuberculosis* lineage 4 shaped by colonial migration and local adaptation." *Science Advances* 4.10, eaat5869.

Carey, AF, JM Rock, IV Krieger, MR Chase, M Fernandez-Suarez, S Gagneux, JC Sacchettini, TR Ioerger, and SM Fortune (2018). "TnSeq of *Mycobacterium tuberculosis* clinical isolates reveals strain-specific antibiotic liabilities." *PLOS Pathogens* 14.3, e1006939.

Casali, N, V Nikolayevskyy, Y Balabanova, SR Harris, O Ignatyeva, I Kontsevaya, J Corander, J Bryant, J Parkhill, S Nejentsev, RD Horstmann, T Brown, and F Drobniewski (2014). "Evolu-

tion and transmission of drug-resistant tuberculosis in a Russian population." *Nature Genetics* 46.3, pp. 279–286.

Castro, RAD, A Ross, L Kamwela, M Reinhard, C Loiseau, J Feldmann, S Borrell, A Trauner, and S Gagneux (2019). "The Genetic Background Modulates the Evolution of Fluoroquinolone-Resistance in *Mycobacterium tuberculosis*." *Molecular Biology and Evolution* msz214.

Chan, ED, V Laurel, MJ Strand, JF Chan, MLN Huynh, M Goble, and MD Iseman (2004). "Treatment and Outcome Analysis of 205 Patients with Multidrug-resistant Tuberculosis." *American Journal of Respiratory and Critical Care Medicine* 169.10, pp. 1103–1109.

Chapman, JR, MRG Taylor, and SJ Boulton (2012). "Playing the End Game: DNA Double-Strand Break Repair Pathway Choice." *Molecular Cell* 47.4, pp. 497–510.

Chaudhuri, S, L Li, M Zimmerman, Y Chen, YX Chen, MN Toosky, M Gardner, M Pan, YY Li, Q Kawaji, JH Zhu, HW Su, AJ Martinot, EJ Rubin, VA Dartois, and B Javid (2018). "Kasugamycin potentiates rifampicin and limits emergence of resistance in *Mycobacterium tuberculosis* by specifically decreasing mycobacterial mistranslation." *eLife* 7. Ed. by M Pai, G Storz, M Pai, and A Diacon, e36782.

Chiang, CY, I Uzoma, RT Moore, M Gilbert, AJ Duplantier, and RG Panchal (2018). "Mitigating the Impact of Antibacterial Drug Resistance through Host-Directed Therapies: Current Progress, Outlook, and Challenges." *mBio* 9.1.

Chien, JY, ST Chien, WY Chiu, CJ Yu, and PR Hsueh (2016). "Moxifloxacin Improves Treatment Outcomes in Patients with Ofloxacin-Resistant Multidrug-Resistant Tuberculosis." *Antimicrobial Agents and Chemotherapy* 60.8, pp. 4708–4716.

Chopra, I, AJ O'Neill, and K Miller (2003). "The role of mutators in the emergence of antibiotic-resistant bacteria." *Drug Resistance Updates* 6.3, pp. 137–145.

Cingolani, P, A Platts, LL Wang, M Coon, T Nguyen, L Wang, SJ Land, X Lu, and DM Ruden (2012). "A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w^{1118}; *iso-2*; *iso-3*." *Fly* 6.2, pp. 80–92.

Cirz, RT, JK Chin, DR Andes, Vd Crécy-Lagard, WA Craig, and FE Romesberg (2005). "Inhibition of Mutation and Combating the Evolution of Antibiotic Resistance." *PLOS Biology* 3.6, e176.

Coeck, N, BC de Jong, M Diels, P de Rijk, E Ardizzoni, A Van Deun, and L Rigouts (2016). "Correlation of different phenotypic drug susceptibility testing methods for four fluoroquinolones in *Mycobacterium tuberculosis*." *Journal of Antimicrobial Chemotherapy* 71.5, pp. 1233–1240.

Cohen, T and M Murray (2004). "Modeling epidemics of multidrug-resistant *M. tuberculosis* of heterogeneous fitness." *Nature Medicine* 10.10, pp. 1117–1121.

Colangeli, R, H Jedrey, S Kim, R Connell, S Ma, UD Chippada Venkata, S Chakravorty, A Gupta, EE Sizemore, L Diem, DR Sherman, A Okwera, R Dietze, WH Boom, JL Johnson, WR Mac Kenzie, and D Alland (2018). "Bacterial Factors That Predict Relapse after Tuberculosis Therapy." *New England Journal of Medicine* 379.9, pp. 823–833.

Cole, ST, R Brosch, J Parkhill, T Garnier, C Churcher, D Harris, SV Gordon, K Eiglmeier, S Gas, CE Barry, F Tekaia, K Badcock, D Basham, D Brown, T Chillingworth, R Connor, R Davies, K Devlin, T Feltwell, S Gentles, N Hamlin, S Holroyd, T Hornsby, K Jagels, A Krogh, J McLean, S Moule, L Murphy, K Oliver, J Osborne, MA Quail, MA Rajandream, J Rogers, S Rutter, K Seeger, J Skelton, R Squares, S Squares, JE Sulston, K Taylor, S Whitehead, and BG Barrell (1998). "Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence." *Nature* 393.6685, pp. 537–544.

Coll, F, R McNerney, JA Guerra-Assunção, JR Glynn, J Perdigão, M Viveiros, I Portugal, A Pain, N Martin, and TG Clark (2014). "A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains." *Nature Communications* 5, p. 4812.

Colman, RE, A Mace, M Seifert, J Hetzel, H Mshaiel, A Suresh, D Lemmer, DM Engelthaler, DG Catanzaro, AG Young, CM Denkinger, and TC Rodwell (2019). "Whole-genome and targeted sequencing of drug-resistant *Mycobacterium tuberculosis* on the iSeq100 and MiSeq: A performance, ease-of-use, and cost evaluation." *PLOS Medicine* 16.4, e1002794.

Comas, I, S Borrell, A Roetzer, G Rose, B Malla, M Kato-Maeda, J Galagan, S Niemann, and S Gagneux (2012). "Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes." *Nature Genetics* 44.1, pp. 106–110.

Comas, I, J Chakravartti, PM Small, J Galagan, S Niemann, K Kremer, JD Ernst, and S Gagneux (2010). "Human T cell epitopes of *Mycobacterium* tuberculosis are evolutionarily hyperconserved". *Nature Genetics* 42.6, pp. 498–503.

Coscolla, M and S Gagneux (2014). "Consequences of genomic diversity in Mycobacterium tuberculosis." *Seminars in Immunology*. Immunity to *Mycobacterium tuberculosis* 26.6, pp. 431–444.

Couce, A, N Alonso-Rodriguez, C Costas, A Oliver, and J Blázquez (2016). "Intrapopulation variability in mutator prevalence among urinary tract infection isolates of *Escherichia coli*." *Clinical Microbiology and Infection* 22.6, 566.e1–566.e7.

Cuevas, JM, A Willemsen, J Hillung, MP Zwart, and SF Elena (2015). "Temporal Dynamics of Intrahost Molecular Evolution for a Plant RNA Virus." *Molecular Biology and Evolution* 32.5, pp. 1132–1147.

Dalton, T, P Cegielski, S Akksilp, L Asencios, JC Caoili, SN Cho, VV Erokhin, J Ershova, MT Gler, BY Kazennyy, HJ Kim, K Kliiman, E Kurbatova, C Kvasnovsky, V Leimane, Mvd Walt, LE Via, GV Volchenkov, MA Yagui, and H Kang (2012). "Prevalence of and risk factors for resistance to second-line drugs in people with multidrug-resistant tuberculosis in eight countries: a prospective cohort study." *The Lancet* 380.9851, pp. 1406–1417.

Dean, AS, H Cox, and M Zignol (2017). "Epidemiology of Drug-Resistant Tuberculosis." *Strain Variation in the Mycobacterium tuberculosis Complex: Its Role in Biology, Epidemiology and Control.*

Ed. by S Gagneux. Advances in Experimental Medicine and Biology. Cham: Springer International Publishing, pp. 209–220.

Decuypere, S, M Vanaerschot, K Brunker, H Imamura, S Müller, B Khanal, S Rijal, JC Dujardin, and GH Coombs (2012). "Molecular Mechanisms of Drug Resistance in Natural *Leishmania* Populations Vary with Genetic Background." *PLOS Neglected Tropical Diseases* 6.2, e1514.

Desai, MM and DS Fisher (2011). "The Balance Between Mutators and Nonmutators in Asexual Populations." *Genetics* 188.4, pp. 997–1014.

Dheda, K, CE Barry, and G Maartens (2016). "Tuberculosis." *The Lancet* 387.10024, pp. 1211–1226.

Dickinson, JM, VR Aber, and DA Mitchison (1977). "Bactericidal Activity of Streptomycin, Isoniazid, Rifampin, Ethambutol, and Pyrazinamide Alone and In Combination Against *Mycobacterium tuberculosis*." *American Review of Respiratory Disease* 116.4, pp. 627–635.

Dippenaar, A, SDC Parsons, SL Sampson, RG van der Merwe, JA Drewe, AM Abdallah, KK Siame, NC Gey van Pittius, PD van Helden, A Pain, and RM Warren (2015). "Whole genome sequence analysis of *Mycobacterium suricattae*." *Tuberculosis* 95.6, pp. 682–688.

Douglas, JT, L Qian, JC Montoya, JM Musser, JDAV Embden, DV Soolingen, and K Kremer (2003). "Characterization of the Manila Family of *Mycobacterium tuberculosis*". *Journal of Clinical Microbiology* 41.6, pp. 2723–2726.

Durão, P, S Trindade, A Sousa, and I Gordo (2015). "Multiple Resistance at No Cost: Rifampicin and Streptomycin a Dangerous Liaison in the Spread of Antibiotic Resistance". *Molecular Biology and Evolution* 32.10, pp. 2675–2680.

Eilertson, B, F Maruri, A Blackman, M Herrera, DC Samuels, and TR Sterling (2014). "High Proportion of Heteroresistance in *gyrA* and *gyrB* in Fluoroquinolone-Resistant *Mycobacterium tuberculosis* Clinical Isolates." *Antimicrobial Agents and Chemotherapy* 58.6, pp. 3270–3275.

Eldholm, V, JHO Pettersson, OB Brynildsrud, A Kitchen, EM Rasmussen, T Lillebaek, JO Rønning, V Crudu, AT Mengshoel, N Debech, K Alfsnes, J Bohlin, CS Pepperell, and F Balloux (2016). "Armed conflict and population displacement as drivers of the evolution and dispersal of *Mycobacterium tuberculosis*." *Proceedings of the National Academy of Sciences* 113.48, pp. 13881–13886.

Ernst, JD (2012). "The immunological life cycle of tuberculosis." *Nature Reviews Immunology* 12.8, pp. 581–591.

Farah, MG, A Tverdal, TW Steen, E Heldal, AB Brantsaeter, and G Bjune (2005). "Treatment outcome of new culture positive pulmonary tuberculosis in Norway." *BMC Public Health* 5.1, p. 14.

Farhat, MR, KR Jacobson, MF Franke, D Kaur, M Murray, and CD Mitnick (2017). "Fluoroquinolone Resistance Mutation Detection Is Equivalent to Culture-Based Drug Sensitivity Testing for Predicting Multidrug-Resistant Tuberculosis Treatment Outcome: A Retrospective Cohort Study." *Clinical Infectious Diseases* 65.8, pp. 1364–1370.

Fenner, L, M Egger, T Bodmer, E Altpeter, M Zwahlen, K Jaton, GE Pfyffer, S Borrell, O Dubuis, T Bruderer, HH Siegrist, H Furrer, A Calmy, J Fehr, JM Stalder, B Ninet, EC Böttger, S Gagneux, and for the Swiss HIV Cohort Study and the Swiss Molecular Epidemiology of Tuberculosis Study Group (2012). "Effect of Mutation and Genetic Background on Drug Resistance in *Mycobacterium tuberculosis*." *Antimicrobial Agents and Chemotherapy* 56.6, pp. 3047–3053.

Feuerriegel, S, HS Cox, N Zarkua, HA Karimovich, K Braker, S Rüsch-Gerdes, and S Niemann (2009). "Sequence Analyses of Just Four Genes To Detect Extensively Drug-Resistant *Mycobacterium tuberculosis* Strains in Multidrug-Resistant Tuberculosis Patients Undergoing Treatment." *Antimicrobial Agents and Chemotherapy* 53.8, pp. 3353–3356.

Finken, M, P Kirschner, A Meier, A Wrede, and EC Böttger (1993). "Molecular basis of streptomycin resistance in *Mycobacterium tuberculosis*: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot." *Molecular Microbiology* 9.6, pp. 1239–1246.

Ford, CB, RR Shah, MK Maeda, S Gagneux, MB Murray, T Cohen, JC Johnston, J Gardy, M Lipsitch, and SM Fortune (2013). "*Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis." *Nature Genetics* 45.7, pp. 784–790.

Franzblau, SG, RS Witzig, JC McLaughlin, P Torres, G Madico, A Hernandez, MT Degnan, MB Cook, VK Quenzer, RM Ferguson, and RH Gilman (1998). "Rapid, Low-Technology MIC Determination with Clinical *Mycobacterium tuberculosis* Isolates by Using the Microplate Alamar Blue Assay." *Journal of Clinical Microbiology* 36.2, pp. 362–366.

Frenoy, A and S Bonhoeffer (2018). "Death and population dynamics affect mutation rate estimates and evolvability under stress in bacteria." *PLOS Biology* 16.5, e2005056.

Fujikawa, H and S Morozumi (2005). "Modeling Surface Growth of *Escherichia coli* on Agar Plates." *Applied and Environmental Microbiology* 71.12, pp. 7920–7926.

Gagneux, S (2018). "Ecology and evolution of *Mycobacterium tuberculosis*." *Nature Reviews Microbiology* 16.4, pp. 202–213.

Gagneux, S, MV Burgos, K DeRiemer, A Enciso, S Muñoz, PC Hopewell, PM Small, and AS Pym (2006a). "Impact of Bacterial Genetics on the Transmission of Isoniazid-Resistant *Mycobacterium tuberculosis*." *PLOS Pathogens* 2.6, e61.

Gagneux, S, K DeRiemer, T Van, M Kato-Maeda, BCd Jong, S Narayanan, M Nicol, S Niemann, K Kremer, MC Gutierrez, M Hilty, PC Hopewell, and PM Small (2006b). "Variable host–pathogen compatibility in *Mycobacterium tuberculosis*." *Proceedings of the National Academy of Sciences* 103.8, pp. 2869–2873.

Gagneux, S, CD Long, PM Small, T Van, GK Schoolnik, and BJM Bohannan (2006c). "The Competitive Cost of Antibiotic Resistance in *Mycobacterium tuberculosis*." *Science* 312.5782, pp. 1944–1946.

Gebrezgabiher, G, G Romha, E Ejeta, G Asebe, E Zemene, and G Ameni (2016). "Treatment Outcome of Tuberculosis Patients under Directly Observed Treatment Short Course and Factors Affecting Outcome in Southern Ethiopia: A Five-Year Retrospective Study." *PLOS ONE* 11.2, e0150560.

Gehre, F, J Otu, K DeRiemer, PFd Sessions, ML Hibberd, W Mulders, T Corrah, BCd Jong, and M Antonio (2013). "Deciphering the Growth Behaviour of *Mycobacterium africanum*." *PLOS Neglected Tropical Diseases* 7.5, e2220.

Gellert, M, K Mizuuchi, MH O'Dea, and HA Nash (1976). "DNA gyrase: an enzyme that introduces superhelical turns into DNA." *Proceedings of the National Academy of Sciences* 73.11, pp. 3872–3876.

Gillespie, SH, AM Crook, TD McHugh, CM Mendel, SK Meredith, SR Murray, F Pappas, PP Phillips, and AJ Nunn (2014). "Four-Month Moxifloxacin-Based Regimens for Drug-Sensitive Tuberculosis." *New England Journal of Medicine* 371.17, pp. 1577–1587.

Giraud, A, I Matic, O Tenaillon, A Clara, M Radman, M Fons, and F Taddei (2001). "Costs and Benefits of High Mutation Rates: Adaptive Evolution of Bacteria in the Mouse Gut." *Science* 291.5513, pp. 2606–2608.

Gong, C, P Bongiorno, A Martins, NC Stephanou, H Zhu, S Shuman, and MS Glickman (2005). "Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C." *Nature Structural & Molecular Biology* 12.4, pp. 304–312.

Gould, CV, PD Sniegowski, M Shchepetov, JP Metlay, and JN Weiser (2007). "Identifying Mutator Phenotypes among Fluoroquinolone-Resistant Strains of *Streptococcus pneumoniae* Using Fluctuation Analysis." *Antimicrobial Agents and Chemotherapy* 51.9, pp. 3225–3229.

Gupta, R, D Barkan, G Redelman-Sidi, S Shuman, and MS Glickman (2011). "Mycobacteria exploit three genetically distinct DNA double-strand break repair pathways." *Molecular Microbiology* 79.2, pp. 316–330.

Gygli, SM, S Borrell, A Trauner, and S Gagneux (2017). "Antimicrobial resistance in *Mycobacterium tuberculosis*: mechanistic and evolutionary perspectives". *FEMS Microbiology Reviews* 41.3, pp. 354–373.

Haldane, JBS (1937). "The Effect of Variation of Fitness." *The American Naturalist* 71.735, pp. 337–349.

Handel, A, E Margolis, and BR Levin (2009). "Exploring the role of the immune response in preventing antibiotic resistance." *Journal of Theoretical Biology* 256.4, pp. 655–662.

Hansen, E, RJ Woods, and AF Read (2017). "How to Use a Chemotherapeutic Agent When Resistance to It Threatens the Patient." *PLOS Biology* 15.2, e2001110.

Hershberg, R, M Lipatov, PM Small, H Sheffer, S Niemann, S Homolka, JC Roach, K Kremer, DA Petrov, MW Feldman, and S Gagneux (2008). "High Functional Diversity in *Mycobacterium tuberculosis* Driven by Genetic Drift and Human Demography." *PLOS Biology* 6.12, e311.

Hershkovitz, I, HD Donoghue, DE Minnikin, GS Besra, OYC Lee, AM Gernaey, E Galili, V Eshed, CL Greenblatt, E Lemma, GK Bar-Gal, and M Spigelman (2008). "Detection and Molecular Characterization of 9000-Year-Old *Mycobacterium tuberculosis* from a Neolithic Settlement in the Eastern Mediterranean." *PLOS ONE* 3.10, e3426.

Hicks, ND, J Yang, X Zhang, B Zhao, YH Grad, L Liu, X Ou, Z Chang, H Xia, Y Zhou, S Wang, J Dong, L Sun, Y Zhu, Y Zhao, Q Jin, and SM Fortune (2018). "Clinically prevalent mutations in *Mycobacterium tuberculosis* alter propionate metabolism and mediate multidrug tolerance." *Nature Microbiology* 3.9, pp. 1032–1042.

Hooper, DC (1999). "Mechanisms of fluoroquinolone resistance." *Drug Resistance Updates* 2.1, pp. 38–55.

Hughes, D and DI Andersson (2017). "Evolutionary Trajectories to Antibiotic Resistance." *Annual Review of Microbiology* 71.1, pp. 579–596.

Huseby, DL, F Pietsch, G Brandis, L Garoff, A Tegehall, and D Hughes (2017). "Mutation Supply and Relative Fitness Shape the Genotypes of Ciprofloxacin-Resistant *Escherichia coli*." *Molecular Biology and Evolution* 34.5, pp. 1029–1039.

Imperial, MZ, P Nahid, PPJ Phillips, GR Davies, K Fielding, D Hanna, D Hermann, RS Wallis, JL Johnson, C Lienhardt, and RM Savic (2018). "A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis." *Nature Medicine* 24.11, pp. 1708–1715.

Jagielski, T, H Ignatowska, Z Bakuła, Ł Dziewit, A Napiórkowska, E Augustynowicz-Kopeć, Z Zwolska, and J Bielecki (2014). "Screening for Streptomycin Resistance-Conferring Mutations in *Mycobacterium tuberculosis* Clinical Isolates from Poland." *PLOS ONE* 9.6, e100078.

Jasenosky, LD, TJ Scriba, WA Hanekom, and AE Goldfeld (2015). "T cells and adaptive immunity to *Mycobacterium tuberculosis* in humans." *Immunological Reviews* 264.1, pp. 74–87.

Javid, B, F Sorrentino, M Toosky, W Zheng, JT Pinkham, N Jain, M Pan, P Deighan, and EJ Rubin (2014). "Mycobacterial mistranslation is necessary and sufficient for rifampicin phenotypic resistance." *Proceedings of the National Academy of Sciences* 111.3, pp. 1132–1137.

Ji, B, N Lounis, C Truffot-Pernot, and J Grosset (1995). "In vitro and in vivo activities of levofloxacin against *Mycobacterium tuberculosis*." *Antimicrobial Agents and Chemotherapy* 39.6, pp. 1341–1344.

Ji, B, N Lounis, C Maslo, C Truffot-Pernot, P Bonnafous, and J Grosset (1998). "In Vitro and In Vivo Activities of Moxifloxacin and Clinafloxacin against *Mycobacterium tuberculosis*." *Antimicrobial Agents and Chemotherapy* 42.8, pp. 2066–2069.

Jindani, A, TS Harrison, AJ Nunn, PP Phillips, GJ Churchyard, S Charalambous, M Hatherill, H Geldenhuys, HM McIlleron, SP Zvada, S Mungofa, NA Shah, S Zizhou, L Magweta, J Shepherd, S Nyirenda, JH van Dijk, HE Clouting, D Coleman, AL Bateson, TD McHugh, PD Butcher, and

DA Mitchison (2014). "High-Dose Rifapentine with Moxifloxacin for Pulmonary Tuberculosis." *New England Journal of Medicine* 371.17, pp. 1599–1608.

Jong, BC de, PC Hill, A Aiken, T Awine, A Martin, IM Adetifa, DJ Jackson-Sillah, A Fox, D Kathryn, S Gagneux, MW Borgdorff, KPWJ McAdam, T Corrah, PM Small, and RA Adegbola (2008). "Progression to Active Tuberculosis, but Not Transmission, Varies by *Mycobacterium tuberculosis* Lineage in The Gambia." *The Journal of Infectious Diseases* 198.7, pp. 1037–1043.

Khan, DD, P Lagerbäck, C Malmberg, AN Kristoffersson, E Wistrand-Yuen, C Sha, O Cars, DI Andersson, D Hughes, EI Nielsen, and LE Friberg (2018). "Predicting mutant selection in competition experiments with ciprofloxacin-exposed *Escherichia coli*." *International Journal of Antimicrobial Agents* 51.3, pp. 399–406.

Kibret, KT, Y Moges, P Memiah, and S Biadgilign (2017). "Treatment outcomes for multidrug-resistant tuberculosis under DOTS-Plus: a systematic review and meta-analysis of published studies." *Infectious Diseases of Poverty* 6.1, p. 7.

Koboldt, DC, Q Zhang, DE Larson, D Shen, MD McLellan, L Lin, CA Miller, ER Mardis, L Ding, and RK Wilson (2012). "VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing." *Genome Research* 22.3, pp. 568–576.

Koch, R (1882). "Die Ätiologie der Tuberkulose." *Berliner klinische Wochenschrift* 15, pp. 221–230.

Laxminarayan, R, A Duse, C Wattal, AKM Zaidi, HFL Wertheim, N Sumpradit, E Vlieghe, GL Hara, IM Gould, H Goossens, C Greko, AD So, M Bigdeli, G Tomson, W Woodhouse, E Ombaka, AQ Peralta, FN Qamar, F Mir, S Kariuki, ZA Bhutta, A Coates, R Bergstrom, GD Wright, ED Brown, and O Cars (2013). "Antibiotic resistance—the need for global solutions." *The Lancet Infectious Diseases* 13.12, pp. 1057–1098.

Lehmann, J (1946). "PARA-AMINOSALICYLIC ACID IN THE TREATMENT OF TUBERCULOSIS." *The Lancet*. Originally published as Volume 1, Issue 6384 247.6384, pp. 15–16.

Leimane, V, G Dravniece, V Riekstina, I Sture, S Kammerer, MP Chen, G Skenders, and TH Holtz (2010). "Treatment outcome of multidrug/extensively drug-resistant tuberculosis in Latvia, 2000–2004." *European Respiratory Journal* 36.3, pp. 584–593.

Levine, C, H Hiasa, and KJ Marians (1998). "DNA gyrase and topoisomerase IV: biochemical activities, physiological roles during chromosome replication, and drug sensitivities." *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1400.1, pp. 29–43.

Levin-Reisman, I, I Ronin, O Gefen, I Braniss, N Shoresh, and NQ Balaban (2017). "Antibiotic tolerance facilitates the evolution of resistance." *Science* 355.6327, pp. 826–830.

Li, H (2011). "A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data." *Bioinformatics* 27.21, pp. 2987–2993.

Li, H, B Handsaker, A Wysoker, T Fennell, J Ruan, N Homer, G Marth, G Abecasis, and R Durbin (2009). "The Sequence Alignment/Map format and SAMtools." *Bioinformatics* 25.16, pp. 2078–2079.

Li, WH (1997). *Molecular Evolution*. Sunderland, Massachusetts, United States: Sinauer Associates Inc.

Lin, PL and JL Flynn (2018). "The End of the Binary Era: Revisiting the Spectrum of Tuberculosis." *The Journal of Immunology* 201.9, pp. 2541–2548.

Lindsey, HA, J Gallie, S Taylor, and B Kerr (2013). "Evolutionary rescue from extinction is contingent on a lower rate of environmental change." *Nature* 494.7438, pp. 463–467.

Liu, CH, H Liu, and B Ge (2017). "Innate immunity in tuberculosis: host defense vs pathogen evasion." *Cellular & Molecular Immunology* 14.12, pp. 963–975.

Loftfield, RB and D Vanderjagt (1972). "The frequency of errors in protein biosynthesis." *Biochemical Journal* 128.5, pp. 1353–1356.

Loiseau, C, F Menardo, A Aseffa, E Hailu, B Gumi, G Ameni, S Berg, L Rigouts, S Robbe-Austerman, J Zinsstag, S Gagneux, and D Brites (2019). "An African origin for *Mycobacterium bovis*." *bioRxiv*, p. 773192.

Luria, SE and M Delbrück (1943). "Mutations of Bacteria from Virus Sensitivity to Virus Resistance." *Genetics* 28.6, pp. 491–511.

Lynch, M, MS Ackerman, JF Gout, H Long, W Sung, WK Thomas, and PL Foster (2016). "Genetic drift, selection and the evolution of the mutation rate." *Nature Reviews Genetics* 17.11, pp. 704–714.

MacGowan, AP (2008). "Clinical implications of antimicrobial resistance for therapy." *Journal of Antimicrobial Chemotherapy* 62.suppl_2, pp. ii105–ii114.

Maciá, MD, D Blanquer, B Togores, J Sauleda, JL Pérez, and A Oliver (2005). "Hypermutation Is a Key Factor in Development of Multiple-Antimicrobial Resistance in *Pseudomonas aeruginosa* Strains Causing Chronic Lung Infections." *Antimicrobial Agents and Chemotherapy* 49.8, pp. 3382–3386.

Maggi, N, CR Pasqualucci, R Ballotta, and P Sensi (1966). "Rifampicin: A New Orally Active Rifamycin." *Chemotherapy* 11.5, pp. 285–292.

Maitre, T, G Petitjean, A Chauffour, C Bernard, N El Helali, V Jarlier, F Reibel, P Chavanet, A Aubry, and N Veziris (2017). "Are moxifloxacin and levofloxacin equally effective to treat XDR tuberculosis?" *Journal of Antimicrobial Chemotherapy* 72.8, pp. 2326–2333.

Malik, M, G Hoatam, K Chavda, RJ Kerns, and K Drlica (2010). "Novel Approach for Comparing the Abilities of Quinolones To Restrict the Emergence of Resistant Mutants during Quinolone Exposure." *Antimicrobial Agents and Chemotherapy* 54.1, pp. 149–156.

Malik, M, A Mustaev, HA Schwanz, G Luan, N Shah, LM Oppegard, EC de Souza, H Hiasa, X Zhao, RJ Kerns, and K Drlica (2016). "Suppression of gyrase-mediated resistance by C7 aryl fluoroquinolones." *Nucleic Acids Res* 44.7, pp. 3304–3316.

Manson, AL, KA Cohen, T Abeel, CA Desjardins, DT Armstrong, CEB Iii, J Brand, TGG Consortium, J Brand, P Jureen, L Malinga, D Nordenberg, AA Velayati, GH Cassell, P Farnia, D Homorodean, MVd Walt, S Hoffner, SB Chapman, SN Cho, A Gabrielian, J Gomez, AM Jodals, M Joloba, P Jureen, JS Lee, L Malinga, M Maiga, D Nordenberg, E Noroc, E Romancenco, A Salazar, W Ssengooba, AA Velayati, K Winglee, A Zalutskaya, LE Via, GH Cassell, SE Dorman, J Ellner, P Farnia, JE Galagan, A Rosenthal, V Crudu, D Homorodean, PR Hsueh, S Narayanan, AS Pym, A Skrahina, S Swaminathan, MVd Walt, D Alland, WR Bishai, T Cohen, S Hoffner, BW Birren, and AM Earl (2017). "Genomic analysis of globally diverse *Mycobacterium tuberculosis* strains provides insights into the emergence and spread of multidrug resistance." *Nature Genetics* 49.3, pp. 395–402.

Maruri, F, TR Sterling, AW Kaiga, A Blackman, YF van der Heijden, C Mayer, E Cambau, and A Aubry (2012). "A systematic review of gyrase mutations associated with fluoroquinolone-resistant *Mycobacterium tuberculosis* and a proposed gyrase numbering system." *Journal of Antimicrobial Chemotherapy* 67.4, pp. 819–831.

Mayer, C and H Takiff (2014). "The Molecular Genetics of Fluoroquinolone Resistance in *Mycobacterium tuberculosis*." *Microbiology Spectrum* 2.4.

McGrath, M, NC Gey van Pittius, PD van Helden, RM Warren, and DF Warner (2014). "Mutation rate and the emergence of drug resistance in *Mycobacterium tuberculosis*." *Journal of Antimicrobial Chemotherapy* 69.2, pp. 292–302.

McKenna, A, M Hanna, E Banks, A Sivachenko, K Cibulskis, A Kernytsky, K Garimella, D Altshuler, S Gabriel, M Daly, and MA DePristo (2010). "The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data." *Genome Research* 20.9, pp. 1297–1303.

McKenzie, D, L Malone, S Kushner, JJ Oleson, and Y SubbaRow (1948). "The effect of nicotinic acid amide on experimental tuberculosis of white mice." *The Journal of Laboratory and Clinical Medicine* 33.10, pp. 1249–1253.

Meehan, CJ, GA Goig, TA Kohl, L Verboven, A Dippenaar, M Ezewudo, MR Farhat, JL Guthrie, K Laukens, P Miotto, B Ofori-Anyinam, V Dreyer, P Supply, A Suresh, C Utpatel, Dv Soolingen, Y Zhou, PM Ashton, D Brites, AM Cabibbe, BCd Jong, Md Vos, F Menardo, S Gagneux, Q Gao, TH Heupink, Q Liu, C Loiseau, L Rigouts, TC Rodwell, E Tagliani, TM Walker, RM Warren, Y Zhao, M Zignol, M Schito, J Gardy, DM Cirillo, S Niemann, I Comas, and AV Rie (2019). "Whole genome sequencing of Mycobacterium tuberculosis : current standards and open issues". *Nature Reviews Microbiology* 17.9, pp. 533–545.

Menardo, F, C Loiseau, D Brites, M Coscolla, SM Gygli, LK Rutaihwa, A Trauner, C Beisel, S Borrell, and S Gagneux (2018). "Treemmer: a tool to reduce large phylogenetic datasets with minimal loss of diversity." *BMC Bioinformatics* 19.1, p. 164.

Merker, M, M Barbier, H Cox, JP Rasigade, S Feuerriegel, TA Kohl, R Diel, S Borrell, S Gagneux, V Nikolayevskyy, S Andres, U Nübel, P Supply, T Wirth, and S Niemann (2018). "Compensatory evolution drives multidrug-resistant tuberculosis in Central Asia." *eLife* 7. Ed. by G Storz, e38200.

Merker, M, C Blin, S Mona, N Duforet-Frebourg, S Lecher, E Willery, MGB Blum, S Rüsch-Gerdes, I Mokrousov, E Aleksic, C Allix-Béguec, A Antierens, E Augustynowicz-Kopeć, M Ballif, F Barletta, HP Beck, CE Barry Iii, M Bonnet, E Borroni, I Campos-Herrero, D Cirillo, H Cox, S Crowe, V Crudu, R Diel, F Drobniewski, M Fauville-Dufaux, S Gagneux, S Ghebremichael, M Hanekom, S Hoffner, Ww Jiao, S Kalon, TA Kohl, I Kontsevaya, T Lillebæk, S Maeda, V Nikolayevskyy, M Rasmussen, N Rastogi, S Samper, E Sanchez-Padilla, B Savic, IC Shamputa, A Shen, LH Sng, P Stakenas, K Toit, F Varaine, D Vukovic, C Wahl, R Warren, P Supply, S Niemann, and T Wirth (2015). "Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage." *Nature Genetics* 47.3, pp. 242–249.

Merle, CS, K Fielding, OB Sow, M Gninafon, MB Lo, T Mthiyane, J Odhiambo, E Amukoye, B Bah, F Kassa, A N'Diaye, R Rustomjee, BC de Jong, J Horton, C Perronne, C Sismanidis, O Lapujade, PL Olliaro, and C Lienhardt (2014). "A Four-Month Gatifloxacin-Containing Regimen for Treating Tuberculosis." *New England Journal of Medicine* 371.17, pp. 1588–1598.

Muller, HJ (1950). "Our load of mutations." *American Journal of Human Genetics* 2.2, pp. 111–176.

Mullis, MN, T Matsui, R Schell, R Foree, and IM Ehrenreich (2018). "The complex underpinnings of genetic background effects." *Nature Communications* 9.1, pp. 1–10.

Munro, SA, SA Lewin, HJ Smith, ME Engel, A Fretheim, and J Volmink (2007). "Patient Adherence to Tuberculosis Treatment: A Systematic Review of Qualitative Research." *PLOS Medicine* 4.7, e238.

Mustaev, A, M Malik, X Zhao, N Kurepina, G Luan, LM Oppegard, H Hiasa, KR Marks, RJ Kerns, JM Berger, and K Drlica (2014). "Fluoroquinolone-Gyrase-DNA Complexes TWO MODES OF DRUG BINDING." *Journal of Biological Chemistry* 289.18, pp. 12300–12312.

Nei, M and WH Li (1979). "Mathematical model for studying genetic variation in terms of restriction endonucleases." *Proceedings of the National Academy of Sciences* 76.10, pp. 5269–5273.

Nerlich, AG, CJ Haas, A Zink, U Szeimies, and HG Hagedorn (1997). "Molecular evidence for tuberculosis in an ancient Egyptian mummy". *The Lancet* 350.9088, p. 1404.

Nhu, NTQ, NTN Lan, NTN Phuong, NvV Chau, J Farrar, and M Caws (2012). "Association of streptomycin resistance mutations with level of drug resistance and *Mycobacterium tuberculosis* genotypes." *The International Journal of Tuberculosis and Lung Disease* 16.4, pp. 527–531.

Nuermberger, EL, T Yoshimatsu, S Tyagi, RJ O'Brien, AN Vernon, RE Chaisson, WR Bishai, and JH Grosset (2004). "Moxifloxacin-containing Regimen Greatly Reduces Time to Culture Con-

version in Murine Tuberculosis." *American Journal of Respiratory and Critical Care Medicine* 169.3, pp. 421–426.

Nunn, AJ, PP Phillips, SK Meredith, CY Chiang, F Conradie, D Dalai, A van Deun, PT Dat, N Lan, I Master, T Mebrahtu, D Meressa, R Moodliar, N Ngubane, K Sanders, SB Squire, G Torrea, B Tsogt, and I Rusen (2019). "A Trial of a Shorter Regimen for Rifampin-Resistant Tuberculosis." *New England Journal of Medicine* 380.13, pp. 1201–1213.

O'Garra, A, PS Redford, FW McNab, CI Bloom, RJ Wilkinson, and MP Berry (2013). "The Immune Response in Tuberculosis." *Annual Review of Immunology* 31.1, pp. 475–527.

Oliver, A and A Mena (2010). "Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance." *Clinical Microbiology and Infection* 16.7, pp. 798–808.

Oliver, A, R Cantón, P Campo, F Baquero, and J Blázquez (2000). "High Frequency of Hypermutable *Pseudomonas aeruginosa* in Cystic Fibrosis Lung Infection". *Science* 288.5469, pp. 1251–1253.

O'Neill, J (2016). *Antimicrobial resistance: tackling a crisis for the health and wealth of nations.* Tech. rep.

Oppong, YEA, J Phelan, J Perdigão, D Machado, A Miranda, I Portugal, M Viveiros, TG Clark, and ML Hibberd (2019). "Genome-wide analysis of *Mycobacterium tuberculosis* polymorphisms reveals lineage-specific associations with drug resistance." *BMC Genomics* 20.1, p. 252.

Örlén, H and D Hughes (2006). "Weak Mutators Can Drive the Evolution of Fluoroquinolone Resistance in *Escherichia coli*." *Antimicrobial Agents and Chemotherapy* 50.10, pp. 3454–3456.

Oxlade, O and M Murray (2012). "Tuberculosis and Poverty: Why Are the Poor at Greater Risk in India?" *PLOS ONE* 7.11, e47533.

Ozbudak, EM, M Thattai, I Kurtser, AD Grossman, and Av Oudenaarden (2002). "Regulation of noise in the expression of a single gene." *Nature Genetics* 31.1, pp. 69–73.

Pandey, B, S Grover, C Tyagi, S Goyal, S Jamal, A Singh, J Kaur, and A Grover (2018). "Dynamics of fluoroquinolones induced resistance in DNA gyrase of *Mycobacterium tuberculosis*." *Journal of Biomolecular Structure and Dynamics* 36.2, pp. 362–375.

Parmar, MM, KS Sachdeva, PK Dewan, K Rade, SA Nair, R Pant, and SD Khaparde (2018). "Unacceptable treatment outcomes and associated factors among India's initial cohorts of multidrug-resistant tuberculosis (MDR-TB) patients under the revised national TB control programme (2007–2011): Evidence leading to policy enhancement." *PLOS ONE* 13.4, e0193903.

Payne, JL, F Menardo, A Trauner, S Borrell, SM Gygli, C Loiseau, S Gagneux, and AR Hall (2019a). "Transition bias influences the evolution of antibiotic resistance in *Mycobacterium tuberculosis*." *PLOS Biology* 17.5, e3000265.

Payne, JL and A Wagner (2019b). "The causes of evolvability and their evolution." *Nature Reviews Genetics* 20.1, pp. 24–38.

Pienaar, E, J Sarathy, B Prideaux, J Dietzold, V Dartois, DE Kirschner, and JJ Linderman (2017). "Comparing efficacies of moxifloxacin, levofloxacin and gatifloxacin in tuberculosis granulo-

mas using a multi-scale systems pharmacology approach." *PLOS Computational Biology* 13.8, e1005650.

Piton, J, S Petrella, M Delarue, G André-Leroux, V Jarlier, A Aubry, and C Mayer (2010). "Structural Insights into the Quinolone Resistance Mechanism of *Mycobacterium tuberculosis* DNA Gyrase." *PLOS ONE* 5.8, e12245.

Portevin, D, S Gagneux, I Comas, and D Young (2011). "Human Macrophage Responses to Clinical Isolates from the *Mycobacterium tuberculosis* Complex Discriminate between Ancient and Modern Lineages." *PLOS Pathogens* 7.3, e1001307.

Pranger, AD, TS van der Werf, JGW Kosterink, and JWC Alffenaar (2019). "The Role of Fluoroquinolones in the Treatment of Tuberculosis in 2019." *Drugs* 79.2, pp. 161–171.

Prunier, AL, B Malbruny, M Laurans, J Brouard, JF Duhamel, and R Leclercq (2003). "High Rate of Macrolide Resistance in *Staphylococcus aureus* Strains from Patients with Cystic Fibrosis Reveals High Proportions of Hypermutable Strains." *The Journal of Infectious Diseases* 187.11, pp. 1709–1716.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rad, ME, P Bifani, C Martin, K Kremer, S Samper, J Rauzier, BN Kreiswirth, J Blazquez, M Jouan, Dv Soolingen, and B Gicquel (2003). "Mutations in Putative Mutator Genes of *Mycobacterium tuberculosis* Strains of the W-Beijing Family." *Emerging Infectious Diseases* 9.7, pp. 838–845.

Raynes, Y, CS Wylie, PD Sniegowski, and DM Weinreich (2018). "Sign of selection on mutation rate modifiers depends on population size." *Proceedings of the National Academy of Sciences* 115.13, pp. 3422–3427.

Read, AF and LH Taylor (2001). "The Ecology of Genetically Diverse Infections." *Science* 292.5519, pp. 1099–1102.

Reha-Krantz, LJ (2010). "DNA polymerase proofreading: Multiple roles maintain genome stability." *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. DNA Polymerase: Structure and Function 1804.5, pp. 1049–1063.

Reiling, N, S Homolka, K Walter, J Brandenburg, L Niwinski, M Ernst, C Herzmann, C Lange, R Diel, S Ehlers, and S Niemann (2013). "Clade-Specific Virulence Patterns of *Mycobacterium tuberculosis* Complex Strains in Human Primary Macrophages and Aerogenically Infected Mice." *mBio* 4.4.

Rigouts, L, N Coeck, M Gumusboga, WB de Rijk, KJM Aung, MA Hossain, K Fissette, HL Rieder, CJ Meehan, BC de Jong, and A Van Deun (2016). "Specific *gyrA* gene mutations predict poor treatment outcome in MDR-TB." *Journal of Antimicrobial Chemotherapy* 71.2, pp. 314–323.

Rigouts, L, P Miotto, M Schats, P Lempens, AM Cabibbe, S Galbiati, V Lampasona, Pd Rijk, DM Cirillo, and BCd Jong (2019). "Fluoroquinolone heteroresistance in *Mycobacterium tuberculosis*: detection by genotypic and phenotypic assays in experimentally mixed populations." *Scientific Reports* 9.1, pp. 1–8.

Robitzek, EH, IJ Selikoff, and GG Ornstein (1952). "Chemotherapy of human tuberculosis with hydrazine derivatives of isonicotinic acid; preliminary report of representative cases." *Quarterly bulletin of Sea View Hospital. New York. Sea View Hospital, Staten Island. Clinical Society* 13.1, pp. 27–51.

Rock, JM, UF Lang, MR Chase, CB Ford, ER Gerrick, R Gawande, M Coscolla, S Gagneux, SM Fortune, and MH Lamers (2015). "DNA replication fidelity in *Mycobacterium tuberculosis* is mediated by an ancestral prokaryotic proofreader." *Nature Genetics* 47.6, pp. 677–681.

Rosche, WA and PL Foster (2000). "Determining Mutation Rates in Bacterial Populations." *Methods* 20.1, pp. 4–17.

Rose, G, T Cortes, I Comas, M Coscolla, S Gagneux, and DB Young (2013). "Mapping of Genotype–Phenotype Diversity among Clinical Isolates of *Mycobacterium tuberculosis* by Sequence-Based Transcriptional Profiling." *Genome Biology and Evolution* 5.10, pp. 1849–1862.

Ruru, Y, M Matasik, A Oktavian, R Senyorita, Y Mirino, LH Tarigan, MJvd Werf, E Tiemersma, and B Alisjahbana (2018). "Factors associated with non-adherence during tuberculosis treatment among patients treated with DOTS strategy in Jayapura, Papua Province, Indonesia." *Global Health Action* 11.1, p. 1510592.

Russell, SL and CM Cavanaugh (2017). "Intrahost Genetic Diversity of Bacterial Symbionts Exhibits Evidence of Mixed Infections and Recombinant Haplotypes". *Molecular Biology and Evolution* 34.11, pp. 2747–2761.

Sarathy, JP, LE Via, D Weiner, L Blanc, H Boshoff, EA Eugenin, CE Barry, and VA Dartois (2018). "Extreme Drug Tolerance of *Mycobacterium tuberculosis* in Caseum." *Antimicrobial Agents and Chemotherapy* 62.2, e02266–17.

Sarkar, R, L Lenders, KA Wilkinson, RJ Wilkinson, and MP Nicol (2012). "Modern Lineages of *Mycobacterium tuberculosis* Exhibit Lineage-Specific Patterns of Growth and Cytokine Induction in Human Monocyte-Derived Macrophages." *PLOS ONE* 7.8, e43170.

Schatz, A, E Bugie, and S Waksman (1944). "Streptomycin, a Substance Exhibiting Antibiotic Activity against Gram-Positive and Gram-Negative Bacteria." *Proceedings of the Society for Experimental Biology and Medicine* 55, pp. 66–69.

Schwarze, K, J Buchanan, JM Fermont, H Dreau, MW Tilley, JM Taylor, P Antoniou, SJL Knight, C Camps, MM Pentony, EM Kvikstad, S Harris, N Popitsch, AT Pagnamenta, A Schuh, JC Taylor, and S Wordsworth (2019). "The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom." *Genetics in Medicine*, pp. 1–10.

Shah, NS, SC Auld, JC Brust, B Mathema, N Ismail, P Moodley, K Mlisana, S Allana, A Campbell, T Mthiyane, N Morris, P Mpangase, H van der Meulen, SV Omar, TS Brown, A Narechania, E Shaskina, T Kapwata, B Kreiswirth, and NR Gandhi (2017). "Transmission of Extensively Drug-Resistant Tuberculosis in South Africa." *New England Journal of Medicine* 376.3, pp. 243–253.

Shitikov, E, S Kolchenko, I Mokrousov, J Bespyatykh, D Ischenko, E Ilina, and V Govorun (2017). "Evolutionary pathway analysis and unified classification of East Asian lineage of *Mycobacterium tuberculosis*." *Scientific Reports* 7.1, pp. 1–10.

Singh, A (2017). "Guardians of the mycobacterial genome: A review on DNA repair systems in Mycobacterium tuberculosis." *Microbiology* 163.12, pp. 1740–1758.

Song, T, Y Park, IC Shamputa, S Seo, SY Lee, HS Jeon, H Choi, M Lee, RJ Glynne, SW Barnes, JR Walker, S Batalov, K Yusim, S Feng, CS Tung, J Theiler, LE Via, HIM Boshoff, KS Murakami, B Korber, CE Barry, and SN Cho (2014). "Fitness costs of rifampicin resistance in *Mycobacterium tuberculosis* are amplified under conditions of nutrient starvation and compensated by mutation in the β subunit of RNA polymerase." *Molecular Microbiology* 91.6, pp. 1106–1119.

Soolingen, Dv, PW Hermans, PEd Haas, DR Soll, and JDv Embden (1991). "Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis." *Journal of Clinical Microbiology* 29.11, pp. 2578–2586.

Spicknall, IH, B Foxman, CF Marrs, and JNS Eisenberg (2013). "A Modeling Framework for the Evolution and Spread of Antibiotic Resistance: Literature Review and Model Categorization." *American Journal of Epidemiology* 178.4, pp. 508–520.

Sreevatsan, S, X Pan, KE Stockbauer, DL Williams, BN Kreiswirth, and JM Musser (1996). "Characterization of *rpsL* and *rrs* mutations in streptomycin-resistant *Mycobacterium tuberculosis* isolates from diverse geographic localities." *Antimicrobial Agents and Chemotherapy* 40.4, pp. 1024–1026.

Staden, R (1996). "The staden sequence analysis package." *Molecular Biotechnology* 5.3, p. 233.

Stamatakis, A (2014). "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies." *Bioinformatics* 30.9, pp. 1312–1313.

Stosic, M, D Vukovic, D Babic, G Antonijevic, KL Foley, I Vujcic, and SS Grujicic (2018). "Risk factors for multidrug-resistant tuberculosis among tuberculosis patients in Serbia: a case-control study". *BMC Public Health* 18.1, p. 1114.

Stucki, D, D Brites, L Jeljeli, M Coscolla, Q Liu, A Trauner, L Fenner, L Rutaihwa, S Borrell, T Luo, Q Gao, M Kato-Maeda, M Ballif, M Egger, R Macedo, H Mardassi, M Moreno, GT Vilanova, J Fyfe, M Globan, J Thomas, F Jamieson, JL Guthrie, A Asante-Poku, D Yeboah-Manu, E Wampande, W Ssengooba, M Joloba, WH Boom, I Basu, J Bower, M Saraiva, SEG Vasconcellos, P Suffys, A Koch, R Wilkinson, L Gail-Bekker, B Malla, SD Ley, HP Beck, BC de Jong, K Toit, E Sanchez-Padilla, M Bonnet, A Gil-Brusola, M Frank, VN Penlap Beng, K Eisenach, I Alani, PW Ndung'u, G Revathi, F Gehre, S Akter, F Ntoumi, L Stewart-Isherwood, NE Ntinginya, A Rachow, M Hoelscher, DM Cirillo, G Skenders, S Hoffner, D Bakonyte, P Stakenas, R Diel, V Crudu, O Moldovan, S Al-Hajoj, L Otero, F Barletta, EJ Carter, L Diero, P Supply, I Comas, S Niemann, and S Gagneux (2016). "*Mycobacterium tuberculosis* lineage 4 comprises globally distributed and geographically restricted sublineages." *Nature Genetics* 48.12, pp. 1535–1543.

Sun, D, K Jeannot, Y Xiao, and CW Knapp (2019). "Editorial: Horizontal Gene Transfer Mediated Bacterial Antibiotic Resistance." *Frontiers in Microbiology* 10.

Sun, H, J Zeng, S Li, P Liang, C Zheng, Y Liu, T Luo, N Rastogi, and Q Sun (2018). "Interaction between *rpsL* and *gyrA* mutations affects the fitness and dual resistance of *Mycobacterium tuberculosis* clinical isolates against streptomycin and fluoroquinolones." *Infection and Drug Resistance*.

Suzuki, T, H Fujita, and JG Choi (2008). "Brief communication: New evidence of tuberculosis from prehistoric Korea—Population movement and early evidence of tuberculosis in far East Asia." *American Journal of Physical Anthropology* 136.3, pp. 357–360.

Swings, T, B Van den Bergh, S Wuyts, E Oeyen, K Voordeckers, KJ Verstrepen, M Fauvart, N Verstraeten, and J Michiels (2017). "Adaptive tuning of mutation rates allows fast response to lethal stress in *Escherichia coli*." *eLife* 6. Ed. by W Shou, e22939.

Taddei, F, M Radman, J Maynard-Smith, B Toupance, PH Gouyon, and B Godelle (1997). "Role of mutator alleles in adaptive evolution." *Nature* 387.6634, pp. 700–702.

Takiff, HE, L Salazar, C Guerrero, W Philipp, WM Huang, B Kreiswirth, ST Cole, WR Jacobs, and A Telenti (1994). "Cloning and nucleotide sequence of *Mycobacterium tuberculosis gyrA* and *gyrB* genes and detection of quinolone resistance mutations." *Antimicrobial Agents and Chemotherapy* 38.4, pp. 773–780.

Takiff, H and E Guerrero (2011). "Current Prospects for the Fluoroquinolones as First-Line Tuberculosis Therapy." *Antimicrobial Agents and Chemotherapy* 55.12, pp. 5421–5429.

Tamura, K, G Stecher, D Peterson, A Filipski, and S Kumar (2013). "MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0." *Molecular Biology and Evolution* 30.12, pp. 2725–2729.

Thomas, JP, CO Baughn, RG Wilkinson, and RG Shepherd (1961). "A New Synthetic Compound with Antituberculous Activity in Mice: Ethambutol (Dextro-2, 2- (Ethylenediimino)-di-1-Butanol)". *American Review of Respiratory Disease* 83.6, pp. 891–893.

Tiberi, S, N du Plessis, G Walzl, MJ Vjecha, M Rao, F Ntoumi, S Mfinanga, N Kapata, P Mwaba, TD McHugh, G Ippolito, GB Migliori, MJ Maeurer, and A Zumla (2018). "Tuberculosis: progress and advances in development of new drugs, treatment regimens, and host-directed therapies." *The Lancet Infectious Diseases* 18.7, e183–e198.

Torres-Barceló, C, G Cabot, A Oliver, A Buckling, and RC MacLean (2013). "A trade-off between oxidative stress resistance and DNA repair plays a role in the evolution of elevated mutation rates in bacteria." *Proceedings of the Royal Society B: Biological Sciences* 280.1757, p. 20130007.

Trauner, A, Q Liu, LE Via, X Liu, X Ruan, L Liang, H Shi, Y Chen, Z Wang, R Liang, W Zhang, W Wei, J Gao, G Sun, D Brites, K England, G Zhang, S Gagneux, CE Barry, and Q Gao (2017). "The within-host population dynamics of *Mycobacterium tuberculosis* vary with treatment efficacy". *Genome Biology* 18.1, p. 71.

Trindade, S, A Sousa, KB Xavier, F Dionisio, MG Ferreira, and I Gordo (2009). "Positive Epistasis Drives the Acquisition of Multidrug Resistance." *PLOS Genetics* 5.7, e1000578.

Tsolaki, AG, AE Hirsh, K DeRiemer, JA Enciso, MZ Wong, M Hannan, YOL Goguet de la Salmoniere, K Aman, M Kato-Maeda, and PM Small (2004). "Functional and evolutionary genomics of *Mycobacterium tuberculosis*: Insights from genomic deletions in 100 strains." *Proceedings of the National Academy of Sciences* 101.14, pp. 4865–4870.

Van Deun, A, AKJ Maug, MAH Salim, PK Das, MR Sarker, P Daru, and HL Rieder (2010). "Short, Highly Effective, and Inexpensive Standardized Treatment of Multidrug-resistant Tuberculosis." *American Journal of Respiratory and Critical Care Medicine* 182.5, pp. 684–692.

Vjecha, MJ, S Tiberi, and A Zumla (2018). "Accelerating the development of therapeutic strategies for drug-resistant tuberculosis." *Nature Reviews Drug Discovery* 17, pp. 607–608.

Vogwill, T, M Kojadinovic, V Furió, and RC MacLean (2014). "Testing the Role of Genetic Background in Parallel Evolution Using the Comparative Experimental Evolution of Antibiotic Resistance". *Molecular Biology and Evolution* 31.12, pp. 3314–3323.

Vogwill, T, M Kojadinovic, and RC MacLean (2016). "Epistasis between antibiotic resistance mutations and genetic background shape the fitness effect of resistance across species of *Pseudomonas*." *Proceedings of the Royal Society B: Biological Sciences* 283.1830, p. 20160151.

Vos, M de, B Müller, S Borrell, PA Black, PD van Helden, RM Warren, S Gagneux, and TC Victor (2013). "Putative Compensatory Mutations in the *rpoC* Gene of Rifampin-Resistant *Mycobacterium tuberculosis* Are Associated with Ongoing Transmission." *Antimicrobial Agents and Chemotherapy* 57.2, pp. 827–832.

Walker, TM, CL Ip, RH Harrell, JT Evans, G Kapatai, MJ Dedicoat, DW Eyre, DJ Wilson, PM Hawkey, DW Crook, J Parkhill, D Harris, AS Walker, R Bowden, P Monk, EG Smith, and TE Peto (2013). "Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study." *The Lancet Infectious Diseases* 13.2, pp. 137–146.

Wan, X, K Koster, L Qian, E Desmond, R Brostrom, S Hou, and JT Douglas (2017). "Genomic analyses of the ancestral Manila family of *Mycobacterium tuberculosis*." *PLOS ONE* 12.4, e0175330.

Wang, S, Y Wang, J Shen, Y Wu, and C Wu (2013). "Polymorphic mutation frequencies in clinical isolates of *Staphylococcus aureus*: the role of weak mutators in the development of fluoroquinolone resistance." *FEMS Microbiology Letters* 341.1, pp. 13–17.

Werngren, J and SE Hoffner (2003). "Drug-Susceptible *Mycobacterium tuberculosis* Beijing Genotype Does Not Develop Mutation-Conferred Resistance to Rifampin at an Elevated Rate." *Journal of Clinical Microbiology* 41.4, pp. 1520–1524.

WHO (2017). *Guidelines for Treatment of Drug-Susceptible Tuberculosis and Patient Care, 2017 Update*. Tech. rep. Geneva, Switzerland.

— (2018). *Global tuberculosis report 2018*. Tech. rep. Geneva, Switzerland.

— (2019a). *Global tuberculosis report 2019*. Tech. rep. Geneva, Switzerland.

— (2019b). *WHO consolidated guidelines on drug-resistant tuberculosis treatment*. Tech. rep. OCLC: 1105238089. Geneva, Switzerland.

Wickham, H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

Wielgoss, S, JE Barrick, O Tenaillon, MJ Wiser, WJ Dittmar, S Cruveiller, B Chane-Woon-Ming, C Médigue, RE Lenski, and D Schneider (2013). "Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load." *Proceedings of the National Academy of Sciences* 110.1, pp. 222–227.

Wiesch, PA zur, R Kouyos, J Engelstädter, RR Regoes, and S Bonhoeffer (2011). "Population biological principles of drug-resistance evolution in infectious diseases". *The Lancet Infectious Diseases* 11.3, pp. 236–247.

Wiesch, PS zur, J Engelstädter, and S Bonhoeffer (2010). "Compensation of Fitness Costs and Reversibility of Antibiotic Resistance Mutations." *Antimicrobial Agents and Chemotherapy* 54.5, pp. 2085–2095.

Winston, CA and K Mitruka (2012). "Treatment Duration for Patients with Drug-Resistant Tuberculosis, United States." *Emerging Infectious Diseases* 18.7, pp. 1201–1202.

Wollenberg, KR, CA Desjardins, A Zalutskaya, V Slodovnikova, AJ Oler, M Quiñones, T Abeel, SB Chapman, M Tartakovsky, A Gabrielian, S Hoffner, A Skrahin, BW Birren, A Rosenthal, A Skrahina, and AM Earl (2017). "Whole-Genome Sequencing of *Mycobacterium tuberculosis* Provides Insight into the Evolution and Genetic Composition of Drug-Resistant Tuberculosis in Belarus." *Journal of Clinical Microbiology* 55.2, pp. 457–469.

Yen, TY, RK Joshi, H Yan, NOL Seto, MM Palcic, and BA Macher (2000). "Characterization of cysteine residues and disulfide bonds in proteins by liquid chromatography/electrospray ionization tandem mass spectrometry". *Journal of Mass Spectrometry* 35.8, pp. 990–1002.

Yew, WW, CK Chan, CC Leung, CH Chau, CM Tam, PC Wong, and J Lee (2003). "Comparative Roles of Levofloxacin and Ofloxacin in the Treatment of Multidrug-Resistant Tuberculosis: Preliminary Results of a Retrospective Study From Hong Kong." *Chest* 124.4, pp. 1476–1481.

Ysern, P, B Clerch, M Castańo, I Gibert, J Barbé, and M Llagostera (1990). "Induction of SOS genes in *Escherichia* coli and mutagenesis in *Salmonella typhimurium* by fluoroquinolones." *Mutagenesis* 5.1, pp. 63–66.

Yu, X, G Wang, S Chen, G Wei, Y Shang, L Dong, T Schön, D Moradigaravand, J Parkhill, SJ Peacock, CU Köser, and H Huang (2016). "Wild-Type and Non-Wild-Type *Mycobacterium tuberculosis* MIC Distributions for the Novel Fluoroquinolone Antofloxacin Compared with Those for Ofloxacin, Levofloxacin, and Moxifloxacin." *Antimicrobial Agents and Chemotherapy* 60.9, pp. 5232–5237.

Zaczek, A, A Brzostek, E Augustynowicz-Kopec, Z Zwolska, and J Dziadek (2009). "Genetic evaluation of relationship between mutations in *rpoB* and resistance of *Mycobacterium tuberculosis* to rifampin." *BMC Microbiology* 9.1, p. 10.

Zhang, L, Q Meng, S Chen, M Zhang, B Chen, B Wu, G Yan, X Wang, and Z Jia (2018). "Treatment outcomes of multidrug-resistant tuberculosis patients in Zhejiang, China, 2009–2013". *Clinical Microbiology and Infection* 24.4, pp. 381–388.

Zhang, X, B Zhao, L Liu, Y Zhu, Y Zhao, and Q Jin (2012). "Subpopulation Analysis of Heteroresistance to Fluoroquinolone in Mycobacterium tuberculosis Isolates from Beijing, China". *Journal of Clinical Microbiology* 50.4, pp. 1471–1474.

Zhang, Y, Y Li, T Li, X Shen, T Zhu, Y Tao, X Li, D Wang, Q Ma, Z Hu, J Liu, J Ruan, J Cai, HY Wang, and X Lu (2019). "Genetic Load and Potential Mutational Meltdown in Cancer Cell Populations". *Molecular Biology and Evolution* 36.3, pp. 541–552.

Zhao, Y, S Xu, L Wang, DP Chin, S Wang, G Jiang, H Xia, Y Zhou, Q Li, X Ou, Y Pang, Y Song, B Zhao, H Zhang, G He, J Guo, and Y Wang (2012). "National Survey of Drug-Resistant Tuberculosis in China". *New England Journal of Medicine* 366.23, pp. 2161–2170.

Zhou, J, Y Dong, X Zhao, S Lee, A Amin, S Ramaswamy, J Domagala, JM Musser, and K Drlica (2000). "Selection of Antibiotic-Resistant Bacterial Mutants: Allelic Diversity among Fluoroquinolone-Resistant Mutations". *The Journal of Infectious Diseases* 182.2, pp. 517–525.

Zignol, M, AS Dean, N Alikhanova, S Andres, AM Cabibbe, DM Cirillo, A Dadu, A Dreyer, M Driesen, C Gilpin, R Hasan, Z Hasan, S Hoffner, A Husain, A Hussain, N Ismail, M Kamal, M Mansjö, L Mvusi, S Niemann, SV Omar, E Qadeer, L Rigouts, S Ruesch-Gerdes, M Schito, M Seyfaddinova, A Skrahina, S Tahseen, WA Wells, YD Mukadi, M Kimerling, K Floyd, K Weyer, and MC Raviglione (2016). "Population-based resistance of Mycobacterium tuberculosis isolates to pyrazinamide and fluoroquinolones: results from a multicountry surveillance project". *The Lancet Infectious Diseases* 16.10, pp. 1185–1192.

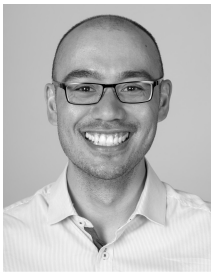
List of Publications

Castro, R.A.D., Ross, A., Kamwela, L., Reinhard, M., Loiseau, C., Feldmann, C., Borrell, S., Trauner, A., and Gagneux, S. "The Genetic Background modulates the Evolution of Fluoroquinolone-Resistance in *Mycobacerium tuberculosis*." *Molecular Biology and Evolution*, 2019.

RHASTIN CASTRO

Address: Margarethenstrasse 59, 4053 Basel, Switzerland
Mobile: +41 78 871 15 11
E-mail: rhastin.castro@alumni.ubc.ca
Skype: rhastin_castro
LinkedIn: https://www.linkedin.com/in/rhastin-castro/

Nationality: Canadian
Swiss Permit: B Permit (eligible for C permit)



Dedicated to improving human health through science, with 6+ years experience researching drug-resistance in tuberculosis.
Strong academic credentials, including a first author publication in a high impact journal and multiple academic honours.
Native English speaker, with excellent interpersonal skills from 8+ years of collaborating with students and associates.
Demonstrated organizational skills from managing and coordinating long-term research projects on infectious diseases.
Capable data analyst, with high proficiency in the R programming language.

CAREER HISTORY

Doctoral Researcher **2015 – March 2020**
Swiss Tropical and Public Health Institute, Basel, Switzerland

- Designed and optimized innovative research projects studying the molecular evolution of drug-resistance in *Mycobacterium tuberculosis*, culminating in research work published in the high-impact journal *Molecular Biology and Evolution*
- Coordinated research teams consisting of an MSc student and laboratory technicians in performing microbiology and molecular biology research in Biosafety Level 1 and Biosafety Level 3 (BSL3) laboratories
- Academically supervised an MSc student for 2 years, coordinated and taught laboratory practical courses for BSc students for 6 semesters, and trained fellow staff members on BSL3 Standard Operating Procedures (SOPs)
- Authored a scientific publication in a high-impact journal, wrote or edited 6 SOPs for use in the BSL3 laboratory, and optimized experimental protocols for internal use
- Detailed organization and management of research projects, including 4 years of experience in planning and implementing long-term *in vitro* experiments that were typically 2 months in duration per experiment
- Analyzed *in vitro*, genomic, and molecular epidemiological data using the R programming language

Residence Advisor **2008 – 2010**
UBC Student Housing and Hospitality Services, Vancouver, Canada

- Cultivated a positive, inclusive, and safe community for student residents through active listening, peer-to-peer coaching, and hosting events, including a touch football (American) league, group dinner outings, and personal health workshops

EDUCATION HISTORY

PhD in Microbiology **2015 – 2019**
Swiss TPH and University of Basel, Basel, Switzerland

- Thesis title: "The Evolution of Fluoroquinolone-Resistance *In Vitro* and in Natural Populations of *Mycobacterium tuberculosis*"
- Honours: *magna cum laude* ("with great praise")

MSc in Infection Biology **2013 – 2015**
Swiss TPH and University of Basel, Basel, Switzerland

- Thesis title: "The Role of Epistatic Interactions in Mycobacterial Drug Resistance Evolution"
- Honours: *magna cum laude* ("with great praise")

BSc in Integrated Sciences **2007 – 2012**
University of British Columbia, Vancouver, Canada

- Focus: Evolutionary Biology, Human pathogens
- Honours: Degree with Distinction, Dean's Honour List, President's Entrance Scholarship

SOFTWARE & LANGUAGES PROFICIENCIES

Microsoft Office	(Expert)	English	(Native Speaker)
R programming language	(Highly Proficient)	German	(Basic, currently enrolled in A2 course)
LaTeX	(Proficient)	Filipino	(Intermediate at listening, basic at speaking)

NOTABLE PRESENTATIONS IN INTERNATIONAL CONFERENCES

Oral Presentations

Joint annual meeting 2017 for the SSM | SSI | SSHH | SSTMP | SSTTM **Aug. 2017**
Basel, Switzerland
Title: "The Fitness and Mutation Rate Estimates of Fluoroquinolone-resistant *Mycobacterium tuberculosis*"

Poster Presentations

EMBO at Basel Life **Sept. 2017 and Sept. 2019**
Basel, Switzerland

Keystone Symposium: New Developments in Our Basic Understanding of Tuberculosis (A5) **Jan. 2017 and Jan. 2019**
Vancouver, Canada

PUBLICATIONS

Castro, R. A. D., Ross, A., Kamwela, L., Reinhard, M., Loiseau, C., Feldmann, J., Borrell, S., Trauner, A, & Gagneux, S. "The Genetic Background modulates the Evolution of Fluoroquinolone-resistance in *Mycobacterium tuberculosis*." (Sept 2019) *Molecular Biology and Evolution*.

Castro, R. A. D., Loiseau, C., Reinhard, M., Feldmann, J., Borrell, S., Trauner, A, & Gagneux, S. "Assessing the Impact of Fluoroquinolone-resistance on the Genetic Diversity in populations of *Mycobacterium tuberculosis*." Manuscript in preparation.

PROFESSIONAL TRAINING

Clinical Research:	"Good Clinical Practice (GCP) for Investigators and Study Teams" by the Swiss TPH	**Dec. 2018**
Leadership:	"Learning How to Lead and to Build a Successful Work Environment" by Carl Emerson, InsideOut Solutions	**Oct. 2018**
	"Successful Team Leadership" by Rob Thomson, RTTA	**Mar. 2018**
Project Management:	"Project Management for Researchers" by Dr. Dimitrije Krstic	**Oct. 2017**
Public Speaking:	"Presentation Training + Video Support" by Sibylle Sommerer, SPEAK! GmbH	**May 2016**

REFERENCES

Prof. Dr. Sebastien Gagneux
Head of the Tuberculosis Research Unit
Head of the Department of Medical Parasitology and Infection Biology at the Swiss TPH
sebastien.gagneux@swisstph.ch
+41 61 284 8369
(Former PhD supervisor)

Prof. Dr. Nino Künzli
Head of the Department of Education and Training at the Swiss TPH
nino.kuenzli@swisstph.ch
+41 61 284 8399
(Reference for community engagement at the Swiss TPH)